

STIC-Biotech/ChemLib

86811

From: Yu, Misook
Sent: Friday, February 14, 2003 9:15 AM
To: STIC-Biotech/ChemLib
Subj ct: 09865879

RECEIVED

FEB 14 2003

COM/CHEM. EN. ISL.
(STIC)

please search SEQ ID NO:1

Examiner Misook Yu, Ph.D.
703-308-2454 (Phone)
Art Unit 1642
CM1-8E18 (Room)
CM1-8E12 (Mail Box)

POINT OF CONTACT:
PAUL SCHULWITZ
TECHNICAL INFO. SPECIALIST
CM1 6B06 TEL. (703) 305-1954

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 2/20
Date Completed: 2/29
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

Yu, Misook

To: STIC-Biotech/ChemLib
Subject: 09865879

please search SEQ ID NO:1

Examiner Misook Yu, Ph.D.
703-308-2454 (Phone)
Art Unit 1642
CM1-8E18 (Room)
CM1-8E12 (Mail Box)

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 23, 2003, 08:00:04 ; Search time 4712 Seconds

(without alignments)
11554.429 Million cell updates/sec

Title: US-09-865-879-1

Perfect score: 1806
Sequence: 1 gggcgatcgttcttctcct.....ggcgagcccgagccgccag 1806

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_da:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: gb_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rtd:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hgtgo_hum:*
40: em_hgtgo_mus:*
41: em_hgtgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1806	100.0	1806	6	AX323409
2	1795	99.4	10884	6	AX409699 Sequence
3	1795	99.4	10884	9	HUM1BP3
4	1685.4	93.3	69887	9	AC091524
5	1141.4	63.2	8700	6	AX345904
6	1021.2	56.5	8700	6	AX345905 Sequence
7	190	10.5	13720	4	AF305712
8	82.6	4.6	107161	2	AC109771
9	78.6	4.4	7218	6	166494
10	74.4	4.1	207420	2	AC078884
11	74.2	4.1	68868	2	AC101770
12	70	3.9	280595	2	AC123860
13	68.2	3.8	187521	2	AC125645
14	68.2	3.8	213985	10	AL591129
15	68.2	3.8	214511	2	AL806529
16	67.8	3.8	112283	10	AL732574
17	67.8	3.8	174099	2	AC109560
18	67.6	3.7	152996	9	AC008177
19	67.6	3.7	176108	2	AC125139
20	67.6	3.7	201893	2	AL833778
21	67.6	3.7	209992	2	AC117827
22	67.4	3.7	177145	2	AC114665
23	67.2	3.7	204431	2	AC096927
24	67	3.7	72358	2	AC101632
25	67	3.7	149211	2	AC102624
26	67	3.7	200885	2	AC119856
27	66.8	3.7	146091	2	AC099260
28	66.8	3.7	162466	2	AC119298
29	66.6	3.7	168073	2	AC129201
30	66.4	3.7	972	11	PM12A126
31	66.4	3.7	150305	2	AC114705
32	66.4	3.7	176107	2	AP002393
33	66.2	3.7	100740	2	AC131216
34	66.2	3.7	193273	10	AC104324
35	66.2	3.7	201650	2	AC119304
36	66.2	3.7	203343	2	AC131133
37	66.2	3.7	242694	2	AC115360
38	66	3.7	147405	2	AC118576
39	66	3.7	168002	2	AC115476
40	66	3.7	181492	2	AC099587
41	66	3.7	201838	2	AC124508
42	65.8	3.6	248201	2	AC115293
43	65.8	3.6	281892	2	AC125486
44	65.6	3.6	88064	2	AC022707
45	65.6	3.6	169709	2	AC105300

ALIGNMENTS

RESULT 1	LOCUS	AX323409	1806 bp	DNA	linear	PAT 07-JAN-2002
	DEFINITION	Sequence 1 from Patent WO0192578.				
	ACCESSION	AX323409				
	VERSION	AX323409.1	GI:18094172			
	KEYWORDS					
	SOURCE	human.				
	ORGANISM	Homo sapiens				
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
		Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
REFERENCE	AUTHORS	Roninson, I.B., Dokmanovic, M. and Chang, B.D.				
	TITLE	Reagents and methods for identifying and modulating expression of genes regulated by retinoids				

Pred. No. is the number of results predicted by chance to have a

JOURNAL Patent: WO 0192578-A 1 06-DEC-2001;
Board of Trustees of the University of Illinois (US)
FEATURES Location/Qualifiers
Source 1. 1806
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 467 a 437 c 517 g 385 t
ORIGIN

Query Match 100.0%; Score 1806; DB 6; Length 1806;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1806; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GGGGATTCGTTTGGTTCCTTCAATTTTCCATGAATCAGAGATCCTGTTGGGTG 60
1 GGGGATTCGTTTGGTTCCTTCAATTTTCCATGAATCAGAGATCCTGTTGGGTG 60
61 CAACGAGATCTAGAGAGAGAGTATCAAGAGAGAGAGAGAGAGAGAGAGAGAG 120
61 CAACGAGATCTAGAGAGAGAGTATCAAGAGAGAGAGAGAGAGAGAGAGAGAG 120
121 ACGGTTTCCTGTAAACAGAGTGTAGTACAGAGAGAGAGAGAGAGAGAGAGAG 180
121 ACGGTTTCCTGTAAACAGAGTGTAGTACAGAGAGAGAGAGAGAGAGAGAGAG 180
181 CATTAAG 240
181 CATTAAG 240
241 TGGTCTCTTTTATTTTGTAGAGTCCAAATTTTATTTTATTTTATTTTATTT 300
241 TGGTCTCTTTTATTTTGTAGAGTCCAAATTTTATTTTATTTTATTTTATTT 300
301 ATTGATCTTTCATTTCTAAAGAGAGATATATATATTAAGAGAGAGAGAGAG 360
301 ATTGATCTTTCATTTCTAAAGAGAGATATATATATTAAGAGAGAGAGAGAG 360
361 GGGGAG 420
361 GGGGAG 420
421 GACAG 480
421 GACAG 480
481 TAACCTTCAACCAAGGCTTGTGCTGGAGAGAGAGAGAGAGAGAGAGAGAG 540
481 TAACCTTCAACCAAGGCTTGTGCTGGAGAGAGAGAGAGAGAGAGAGAGAG 540
541 TACTCCAAACTGCAAAAGGCTCTTCAAGTACCTATCCACTCTCAAGCAAGCT 600
541 TACTCCAAACTGCAAAAGGCTCTTCAAGTACCTATCCACTCTCAAGCAAGCT 600
601 ATTCAACGAGGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
601 ATTCAACGAGGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
661 CCCACAAACTCTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
661 CCCACAAACTCTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
721 CCAGGCTCTGTGTGTGTCACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
721 CCAGGCTCTGTGTGTGTCACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
781 GTGCCAG 840
781 GTGCCAG 840
841 TACTGAG 900
841 TACTGAG 900

901 AATAGGAG 960
901 AATAGGAG 960
961 AATTAGCTTTCAG 1020
961 AATTAGCTTTCAG 1020
1021 CTCATTTGAAAG 1080
1021 CTCATTTGAAAG 1080
1081 AATATTTGAAAG 1140
1081 AATATTTGAAAG 1140
1141 CTTCGGGTGTAG 1200
1141 CTTCGGGTGTAG 1200
1201 AGGCCCCGGGAG 1260
1201 AGGCCCCGGGAG 1260
1261 GGGGTATATTTAAACCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
1261 GGGGTATATTTAAACCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
1321 GTCCCAAG 1380
1321 GTCCCAAG 1380
1381 AATGCGAG 1440
1381 AATGCGAG 1440
1441 CGAG 1500
1441 CGAG 1500
1501 GAGTGAAG 1560
1501 GAGTGAAG 1560
1561 CTTGGGTTCTTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
1561 CTTGGGTTCTTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
1621 CTTGGGTTCTTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
1621 CTTGGGTTCTTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
1681 GGGGCGGAG 1740
1681 GGGGCGGAG 1740
1741 GGTTCGGGAG 1800
1741 GGTTCGGGAG 1800
1801 GCCCAG 1806
1801 GCCCAG 1806
1801 GCCCAG 1806

RESULT 2
AX409699 10884 bp DNA linear PAT 14-JUN-2002
LOCUS AX409699
DEFINITION Sequence 2346 from Patent WO0229103.
ACCESSION AX409699
VERSION AX409699.1 GI:21442404
KEYWORDS

RESULT 3
LOCUS HUMIBP3 10884 bp DNA linear PRI 04-APR-2002
DEFINITION Homo sapiens growth factor-binding protein-3 precursor (IGFBP3)
ACCESSION M35878 J05537 J05538 M35879 M35880 M35881 M35882 M35883 M35884
M35878.1 GI:184522
VERSION M35878.1
KEYWORDS Insulin-like growth factor binding protein 3.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 10884)
AUTHORS Cubbage,M.L., Suwanichkul,A. and Powell,D.R.
TITLE Insulin-like growth factor binding protein-3. Organization of the
human chromosomal gene and demonstration of promoter activity
JOURNAL J. Biol. Chem. 265 (21), 12642-12649 (1990)
MEDLINE 90324259
PubMed 1695633
COMMENT On Apr 4, 2002 this sequence version replaced gi:763445.
Draft entry and computer-readable sequence kindly submitted by
D.R.Powell, 03-JUL-1990.
FEATURES
Source Location/Qualifiers
1..10884
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HL1006d"
/cell_type="Leukocyte"
102..10775
/gene="IGFBP3"
102..1905
/gene="IGFBP3"
/note="G00-126-724"
1808..1821
/gene="IGFBP3"
/note="G00-126-724"
1876..1881
/bound_molety="Spl and Ab-2"
/gene="IGFBP3"
/note="G00-126-724"
1906..2440
/note="G00-126-724"
join(1906..2440,5726..5952,6497..6616,8212..10775)
/gene="IGFBP3"
/product="growth factor-binding protein-3"
exon
1906..2440
/gene="IGFBP3"
/note="G00-126-724"
/number=1
join(2038..2440,5726..5952,6497..6616,8212..8337)
/gene="IGFBP3"
/note="The AAs encoded by bases 5728-5736 and 5917-5926
may be Asn-linked glycosylation sites; Insulin-like
precursor"
/codon_start=1
/product="growth factor-binding protein-3 precursor"
/protein_id="AAAS2706.1"
/db_xref="GI:386791"
/db_xref="GDB:G00-126-724"
/translation="MQRARPTMAALTLVLRGPVVARAGASGGAGPYVRCPCD
ARALACQAPPAVCAELVREPCGCCCLTALSEQPCGTYPRCCSGLRQCSPPRAR
PLQALDGRGICLVNASVSRRLRAYLLPAPRAGNASSESEDSAGSVSPSSYHRV
SDPKFHLPLHSK111IKGHANDSQRYKVDYESQYDTQNFSSSESRRETEYGCRRME
DTLNLKFLNLVSPRGVHIPNDKKGFYKKQCRPSKGRKRGFCWCVDKYGQPLPGYT
TKREDVHCYSNMQSK"
2038..2118
/gene="IGFBP3"
/note="G00-126-724"
join(2119..2440,5726..5952,6497..6616,8212..8334)
/gene="IGFBP3"
/product="growth factor-binding protein-3"
5726..5952
/gene="IGFBP3"
exon
5726..5952
/gene="IGFBP3"

exon
/note="G00-126-724"
/number=2
6497..6616
/gene="IGFBP3"
/note="G00-126-724"
/number=3
8212..10775
/gene="IGFBP3"
/note="G00-126-724"
/number=4
10751..10756
/gene="IGFBP3"
/note="G00-126-724"
polyA_signal
BASE COUNT 2796 a 2578 c 2737 g 2773 t
ORIGIN
Query Match 99.4%; Score 1795; DB 9; Length 10884;
Best Local Similarity 99.9%; Pred. No. 0; Mismatches 0; Indels 1; Gaps 1;
Matches 1806; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 GGGGATTCGTTTGTTCCTTCATATTTCCAAATGAAATCAGAGATCTGTTGGTGT 60
DB 101 GGGGATTCGTTTGTTCCTTCATATTTCCAAATGAAATCAGAGATCTGTTGGTGT 160
QY 61 CAACGAGATCTAGAGAGAGGTGATACAGAAAGC-AACGCAAGCGGATTTATG 119
DB 161 CAACGAGATCTAGAGAGAGGTGATACAGAAAGCAACGCAAGCGGATTTATG 220
QY 120 CACGTTTCTGTAAACAAGGTTGAGTACGACGCTGACCTGACCTGAGGAGAAC 179
DB 221 CACGTTTCTGTAAACAAGGTTGAGTACGACGCTGACCTGAGGAGAAC 280
QY 180 TCATAGAAATAGCGTCTGGGCTTGTACCCGGGGCCCTCAATGTTCTTGTCT 239
DB 281 TCATAGAAATAGCGTCTGGGCTTGTACCCGGGGCCCTCAATGTTCTTGTCT 340
QY 240 TTGGCTCTTTTATTGTAGAGGTCCAAATTTATTTATTTATTTATTTATTTAT 299
DB 341 TTGGCTCTTTTATTGTAGAGGTCCAAATTTATTTATTTATTTATTTATTTAT 400
QY 300 AATTGATCTTCATCTAAAGAGAGATATATGTTAAAGAGAGCTGATGATAT 359
DB 401 AATTGATCTTCATCTAAAGAGAGATATATGTTAAAGAGAGCTGATGATAT 460
QY 360 GGGGAGAGAGGTGACAGGGGGGAAAGGAGAGAGAGAGAGAGAGAGAGAGAG 419
DB 461 GGGGAGAGAGGTGACAGGGGGGAAAGGAGAGAGAGAGAGAGAGAGAGAGAG 520
QY 420 GGAACAAGAGAGACTGGGCGAGAGATGATTAGAGAGAGAGAGAGAGAGAGAG 479
DB 521 GGAACAAGAGAGACTGGGCGAGAGATGATTAGAGAGAGAGAGAGAGAGAGAG 580
QY 480 TTAACTTCAACCAAGGCTCGTCGTGAGGGGAATGAGGAGCTCGATTGCTATTA 539
DB 581 TTAACTTCAACCAAGGCTCGTCGTGAGGGGAATGAGGAGCTCGATTGCTATTA 640
QY 540 CTACTCCAAACTGCAAGAGGCTCTCAAGTCACTATCCACTGCAAGCAAGCTGC 599
DB 641 CTACTCCAAACTGCAAGAGGCTCTCAAGTCACTATCCACTGCAAGCAAGCTGC 700
QY 600 AATTTCACAGCGTTCAGAAAGTCTCCCGGAGGCTTCACCGTTCCTCCACTCAG 659
DB 701 AATTTCACAGCGTTCAGAAAGTCTCCCGGAGGCTTCACCGTTCCTCCACTCAG 760
QY 660 CCCCAAAACTCTTGGAAAAGTGCCTTAATAATTTATCTCAATCTCTGAGCC 719
DB 761 CCCCAAAACTCTTGGAAAAGTGCCTTGAATAATTTATCTCAATCTCTGAGCC 820
QY 720 ACCAGGCTCTCTGTGTGTCACCAAGAGAGGGGTGCGAGACAAATGGAAGAACTCG 779
DB 821 ACCAGGCTCTCTGTGTGTCACCAAGAGAGGGGTGCGAGACAAATGGAAGAACTCG 880
QY 780 AGTGCAGAGAAAGCCGACAGAGATTACAGCACTCAGCGGCAATTCGCCCGCAACT 839

Db	881	AGTCCGAGAGAAAGCCGACAGAGATTACACGACCTCAGCGCCCAATTGGCCCCGAACT	940
Oy	840	TTACGAAAAGTGTGTTGATTTGCAGCATATAGCTAGAAATCCAAAGCATCGAATAATACG	899
Db	941	TTACTGAAGATGTTTAGATTGCAGAAATATAGCTAGATCCCAAGCATGAGAAATACAG	1000
Oy	900	TAAATACGAAGTCGCTTCAAAAAATGACAAATGAAAATTGCTTTTAAAGGACTATTGGT	959
Db	1001	TAAATACGAAGTCGCTTCAAAAAATGACAAATGAAAATTGCTTTTAAAGGACTATTGGT	1066
Oy	960	TAAATAGCTTCACACAGTGGCCAGTTATATGCTTATATTTCTTTTGTCGTGGGGTAA	101
Db	1061	TAAATAGCTTCACACAGTGGCCAGTTATATGCTTATATTTCTTTTGTCGTGGGGTAA	112
Oy	1020	ACTGCATTGTAAAAACATATATCAGGAGCAATACCCAAAGCAAGAAGAACATTTGCTATTAA	107
Db	1121	ACTGCATTGTAAAAACATATATCAGGAGCAATATCCCAAGACAAAGAAAGAACATTTGCTATTAA	118
Oy	1080	AAATATTGTGAAGAACCCCTGGCCTTAAGAGAGATTCGTGGCGGTCCACACTTAATTAGGGG	113
Db	1181	AAATATTGTGAAGAACCCCTGGCCTTAAGAGAGATTCGTGGCGGTCCACACTTAATTAGGGG	124
Oy	1140	ACTTGGGCTGATGACACAGCTGAGAGCTCTTTCGCGTTGGAAGTAAGCCTGGAAAGGCG	119
Db	1241	ACTTGGGCTGATGACACAGCTGAGAGCTCTTTCGCGTTGGAAGTAAGCCTGGAAAGGCG	130
Oy	1200	AAGGCCCCGGGGGCATTCCTTCAGATGCGTAATTTGGGGCCCCCTGGGGATATTAACAAAGGCCAG	125
Db	1301	AAGGCCCCGGGGGCATTCCTTCAGATGCGTAATTTGGGGCCCCCTGGGGATATTAACAAAGGCCAG	136
Oy	1260	CGGGTGTAAATTTAAACCCGCGAGTGCCCTTGTCCTCCGAGACCCCAATGTAACTAGAAA	131

AC091524/C	LOCUS	AC091524	69887 bp	DNA	linear	PRI 07-NOV-2001
	DEFINITION	Homo sapiens BAC clone RP11-132L11	from 7,	complete	sequence.	
	ACCESSION	AC091524				
	VERSION	AC091524.4	GI:16756331			
	KEYWORDS	HTG.				
	SOURCE	Homo sapiens.				
	ORGANISM	Homo sapiens				
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	AUTHORS	1 (bases 1 to 69887)				
	TITLE	Sulston, J.E. and Waterston, R.				
	JOURNAL	Toward a complete human genome sequence				
	MEDLINE	Genome Res. 8 (11), 1097-1108 (1998)				
	PUBMED	99063792				
	PUBMED	9847074				
REFERENCE	AUTHORS	2 (bases 1 to 69887)				
	TITLE	Milvanev, E. and Haakenson, W.				
	JOURNAL	The sequence of Homo sapiens BAC clone RP11-132L11				
	REFERENCE	Unpublished				
	AUTHORS	3 (bases 1 to 69887)				
	TITLE	Waterston, R.H.				
	JOURNAL	Direct Submission				
	REFERENCE	Submitted (29-APR-2001) Genome Sequencing Center, Washington				
	AUTHORS	University School of Medicine, 4444 Forest Park Parkway, St. Louis,				
	TITLE	MO 63108, USA				
	JOURNAL	4 (bases 1 to 69887)				
REFERENCE	AUTHORS	Waterston, R.				
	TITLE	Direct Submission				
	JOURNAL	Submitted (07-NOV-2001) Department of Genetics, Washington				
	REFERENCE	University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA				
	AUTHORS	On Nov 7, 2001 this sequence version replaced gi:14029081.				
COMMENT						

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30): an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION: The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, director), John D. McComb in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRN/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION: The RPc1-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.-Y., Zhao, B., Frangén, E., Toton, M., Catanesse, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pletzer de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBACE3.6
NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is Rp11-63818; the clone sequenced to the right is Rp11-571N3, 200 bp overlap. Actual start of this clone is at base position 1 of Rp11-132L11; actual end is at base position 120387 of Rp11-571N3.

There are polymorphic base pair differences in the overlap between the clone Rp11-132L11, Rp11-63818, Rp11-571N3 and redundant clone Rp11-14802.

```
FEATURES
Source
    Location/Qualifiers
        1..69887
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="7"
            /map="7"
            /clone="Rp11-132L11"
            /clone_id="RPI1-11"
            367..464
            /rpt_family="MER1_type"
            1326..1440
            /rpt_family="MIR"
            2228..2260
            /rpt_family="(CA)n"
            2462..2467
            /note="match to EST AA368830 (NID:g2021148)"
            2913..3072
            /rpt_family="CT-rich"
            3047..3354
            /rpt_family="Alu"
            4681..4752
            /rpt_family="L2"
            5158..5180
            /rpt_family="AT-rich"
            5161..5462
            /rpt_family="Alu"
            5670..5887
            /rpt_family="L1"
            6217..6356
            /rpt_family="MIR"
            7456..7958
            /note="match to EST BG695859 (NID:g13960406)"
            7461..7958
            /note="match to EST AL551522 (NID:g12889551)"
            7462..7958
            /note="similar to Homo sapiens EST AL553038 (NID:g12892497)"
            7462..7825
            /note="match to EST AL551976 (NID:g12890432)"
            7463..7958
            /note="match to EST AL553501 (NID:g12893394)"
            7467..7919
            /note="match to EST AL570357 (NID:g12926584)"
            7470..7958
            /note="similar to Homo sapiens EST AL552017 (NID:g12890512)"
            7486..7958
            /note="match to EST AL531252 (NID:g12794745)"
            7548..7924
            /note="similar to Bos taurus EST BG666453 (NID:g10027044)"
            7802..7958
            /note="similar to Mus musculus EST A1892189 (NID:g5598091)"
            7807..7958
            /note="similar to Homo sapiens EST AL545650 (NID:g12878172)"
            7812..7958
            /note="match to EST AV650862 (NID:g9871876)"
            7816..7958
            /note="match to EST AL570706 (NID:g12927278)"
            7845..7958
            /note="similar to Homo sapiens EST AL571370 (NID:g12928599)"
```

```
misc_feature 7849..7958
             /note="match to EST AL577394 (NID:g12940479)"
misc_feature 7862..7958
             /note="match to EST AA368988 (NID:g2021308)"
misc_feature 7880..7958
             /note="similar to Homo sapiens EST AL571308 (NID:g12928474)"
misc_feature 7931..7958
             /note="match to EST AL571695 (NID:g12929247)"
repeat_region 7969..8031
             /rpt_family="C-rich"
             7985..7990
             /note="match to EST AL569218 (NID:g12924337)"
             8513..8676
             /note="match to EST AA077843 (NID:g1837317)"
             8551..8570
             /note="match to EST BG422582 (NID:g13329088)"
             8596..9243
             /note="match to EST AV685060 (NID:g10286923)"
             8596..9195
             /note="match to EST AV692462 (NID:g10294325)"
             8807..9098
             /note="match to EST AV687868 (NID:g10289731)"
             9103..9369
             /note="match to EST AV659647 (NID:g9880661)"
             9154..9241
             /note="match to EST AV687868 (NID:g10289731)"
             9155..9156
             /note="match to EST W01199 (NID:g1272178) za59a07.r1"
             9172..9457
             /note="match to EST T46868 (NID:g648854) yb94d09.r1"
             9296..9718
             /note="match to EST AW962557 (NID:g8152393)"
             9424..9519
             /note="match to EST AV659647 (NID:g9880661)"
             9473..9674
             /note="similar to Mus musculus EST A1892189 (NID:g5598091)"
             m188b01.y1
             9505..9906
             /note="match to EST AL531252 (NID:g12794745)"
             9505..9674
             /note="match to EST AL551522 (NID:g12889551)"
             9505..9674
             /note="match to EST AL553501 (NID:g12893394)"
             9505..9674
             /note="match to EST AL570706 (NID:g12927278)"
             9505..9674
             /note="match to EST AL571695 (NID:g12929247)"
             9505..9674
             /note="match to EST AL577394 (NID:g12940479)"
             9505..9674
             /note="match to EST AV650862 (NID:g9871876)"
             9505..9674
             /note="match to EST BG695859 (NID:g13960406)"
             9505..9674
             /note="match to EST BG695859 (NID:g13960406)"

Query Match 93.3% Score 1685.4 DB 9 Length 69887;
Best Local Similarity 98.7% Pred. No. 0;
Matches 1794; Conservative 0; Mismatches 11; Indels 12; Gaps 9;

OY 1 GGGGATTCGTTTGGTTCCTTCATTTTCCAAATGAAATCAGAGATCTGTTCTGGGTGT 60
    |||||
DB 42043 GGGGATTCGTTTGGTTCCTTCATTTTCCAAATGAAATCAGAGATCTGTTCTGGGTGT 41984
    |||||

OY 61 CAACGAGATCTAGAGAGAGTGTATCAAGGAAAG-AACAGCAGCAGCGATTATGG 119
    |||||
DB 41983 CAACGAGATCTAGAGAGAGTGTATCAAGGAAAG-AACAGCAGCAGCGATTATGG 41924
    |||||

OY 120 CACGGTTTCCTTAACAAGGTTGAGTGTAGCCACAGCCTGAGCAGCTGGGAGAGAGC 179
    |||||
DB 41923 CACGGTTTCCTTAACAAGGTTGAGTGTAGCCACAGCCTGAGCAGCTGGGAGAGAGC 41864
    |||||

OY 180 TCATAGAAATGACGCTGCTGGGCTTCGTCACCCCGGGGCTTCATTTGTTCTGTCT 239
```

||||| 41863 TCATAGAAAAAGACGGTCTGGCTTCGTACCCCGGGGACATCATGTTCTTGCT 41804
 QY 240 TTGGCTCTTTTATTTAGAGTCAATTTATTTATTTAGTACAGAGGAGCA 299
 Db 41803 TTGGCTCTTTTATTTAGAGTCAATTTATTTATTTAGTACAGAGGAGCA 41744
 QY 300 AATTGATCTTTCCATTTTAAAGAGAGTATATGTATATTAAGAGAGCTATATAT 359
 Db 41743 AATTGATCTTTCCATTTTAAAGAGAGTATATGTATATTAAGAGAGCTATATAT 41684
 QY 360 GGGGGAAGAGTGGACAGGGGGGAAAAAGGAGAGACGAGAGAGAAAGGAGAGAG 419
 Db 41683 GGGGGAAGAGTGGACAGGGGGGAAAAAGGAGAGACGAGAGAGAAAGGAGAGAG 41624
 QY 420 GCACAAAGAGAGACACTGGGCGAGAGATGATAGAGAGAGAGAAATATGATGATAA 479
 Db 41623 GCACAAAGAGAGACACTGGGCGAGAGATGATAGAGAGAGAGAAATATGATGATAA 41564
 QY 480 TTAACTTCAACCAAGGCTTCTGCTCGAGAGGGAATGAGAGAGCTCTGATTTGCTATTA 539
 Db 41563 TTAACTTCAACCAAGGCTTCTGCTCGAGAGGGAATGAGAGAGCTCTGATTTGCTATTA 41505
 QY 540 CTAACTTCAACCAAGGCTTCTGCTCGAGAGGGAATGAGAGAGCTCTGATTTGCTATTA 599
 Db 41504 CTAACTTCAACCAAGGCTTCTGCTCGAGAGGGAATGAGAGAGCTCTGATTTGCTATTA 41445
 QY 600 AATTTCACAGGCTTCAGAGAAAGTCTCTCCGCGAGGCTTCACCGTCCACATCCAC 659
 Db 41444 AATTTCACAGGCTTCAGAGAAAGTCTCTCCGCGAGGCTTCACCGTCCACATCCAC 41385
 QY 660 CCCCAACAACTCTTGGAAAAAGTCCCTCCGCGAGGCTTCACCGTCCACATCCAC 719
 Db 41384 CCCCAACAACTCTTGGAAAAAGTCCCTCCGCGAGGCTTCACCGTCCACATCCAC 41325
 QY 720 ACCAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 778
 Db 41324 ACCAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 41265
 QY 779 GAGTCCAGAGAGAGCCGAGAGAGTACAGGAGCTCAGCGAGCTGCGAATTCG -CCCGGAA 837
 Db 41264 GAGTCCAGAGAGAGCCGAGAGAGTACAGGAGCTCAGCGAGCTGCGAATTCG -CCCGGAA 41205
 QY 838 CTTTACTGAAAAAGTGTATGATTCAGAGATATAGCTAGATCCCAAGCATGAGAAATAC 897
 Db 41204 CTTTACTGAAAAAGTGTATGATTCAGAGATATAGCTAGATCCCAAGCATGAGAAATAC 41145
 QY 898 AGTAATAGCAAGTGGCTCAAAAAATGCAATGCAAAATGGCTATTAAGGACTATTG 957
 Db 41144 AGTAATAGCAAGTGGCTCAAAAAATGCAATGCAAAATGGCTATTAAGGACTATTG 41085
 QY 958 GTTAATACGTTTACAGAGTCCAGTATTTGCTTTATTTATTTGCTGAGGCTG 1017
 Db 41084 GTTAATACGTTTACAGAGTCCAGTATTTGCTTTATTTATTTGCTGAGGCTG 41025
 QY 1018 AAACCTCATTGAAAAATATATGAGGAGATATCCCAAGCAAGAAAGAGTGTCAAT 1077
 Db 41024 AAACCTCATTGAAAAATATATGAGGAGATATCCCAAGCAAGAAAGAGTGTCAAT 40965
 QY 1078 TAAATATTTGAAAAAGCCCTGCTTAAGA -GCATTCGCTGCGGCTGCACTTTAATG 1136
 Db 40964 TAAATATTTGAAAAAGCCCTGCTTAAGA -GCATTCGCTGCGGCTGCACTTTAATG 40905
 QY 1137 GGGACTTGGGCTGAGCAACAGTGAAGTCTTCTGCTGAGAGTAAAGCTGGAAG 1196
 Db 40904 GAGACTTGGGCTGAGCAACAGTGAAGTCTTCTGCTGAGAGTAAAGCTGGAAG 40845
 QY 1197 GCG -- AAGCCCGCGGCGATCTTCAAGTGCATTTGTTGGCCCTGGGAGATATAACA 1253
 Db 40844 GGGCGAGAGCCCGCGGCGATCTTCAAGTGCATTTGTTGGCCCTGGGAGATATAACA 40785
 QY 1254 GCGCAGCGGCTGTAATTAACCCGCAATGCTTGGCTCCGAGAACCCAAATGTAAAT 1313
 Db 1313 GCGCAGCGGCTGTAATTAACCCGCAATGCTTGGCTCCGAGAACCCAAATGTAAAT

Db 40784 GCGCAGCGGCTGTAATTAACCCGCAATGCTTGGCTCCGAGAACCCAAATGTAAAT 40725
 QY 1314 CAGAAATGTCACCAAGCTTCGCTGCGCAACGAAATTAATTTAGAAAGCTCCACAGAGT 1373
 Db 40724 CAGAAATGTCACCAAGCTTCGCTGCGCAACGAAATTAATTTAGAAAGCTCCACAGAGT 40665
 QY 1374 ACACAGAAATGCGGAGCGCTGTATGCCAGTTTCCCGACACCGGCTGCGCAGGGAGAC 1433
 Db 40664 ACACAGAAATGCGGAGCGCTGTATGCCAGTTTCCCGACACCGGCTGCGCAGGGAGAC 40605
 QY 1434 CTCACCCCGAGACCGGAGGAGGTAAGGCGGCGGGGCAAGAGATCGGGGGTGTAGT 1493
 Db 40604 CTCACCCCGAGACCGGAGGAGGTAAGGCGGCGGGGCAAGAGATCGGGGGTGTAGT 40545
 QY 1494 TGGCCAGAGTGAAGTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1553
 Db 40544 TGGCCAGAGTGAAGTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 40485
 QY 1554 GGGCACAACCTTGGTCTTGTAGACGACAAAGTGA -CGGGCTCCGGGCTGAGCAGAGG 1611
 Db 40484 GGGCACAACCTTGGTCTTGTAGACGACAAAGTGA -CGGGCTCCGGGCTGAGCAGAGG 40425
 QY 1612 AGCAGGTGCGCGGCGAGTGTGAGGTGCACGCGCCCGAGCTGGGCGCGGCTGCTAGG 1671
 Db 40424 AGCAGGTGCGCGGCGAGTGTGAGGTGCACGCGCCCGAGCTGGGCGCGGCTGCTAGG 40365
 QY 1672 GCGAAGCAGAGG -CCCGGAGCGCTGCTGCGCGAGACCGCGCCCTCCCAACCCCACT 1730
 Db 40364 GCGAAGCAGAGG -CCCGGAGCGCTGCTGCGCGAGACCGCGCCCTCCCAACCCCACT 40305
 QY 1731 CTTGGGCGCG -CGTTCCGGGCGCTGCTGCGCGACCGCGGCTTCTATATAGGCGCGC 1789
 Db 40304 CTTGGGCGCG -CGTTCCGGGCGCTGCTGCGCGACCGCGGCTTCTATATAGGCGCGC 40245
 QY 1790 GCGCGCGGCGCGCGCAG 1806
 Db 40244 GCGCGCGGCGCGCGCAG 40228

RESULT 5
 AX345904 8700 bp DNA linear PAT 01-FEB-2002
 LOCUS AX345904
 DEFINITION Sequence 975 from Patent WO0200928.
 ACCESSION AX345904
 VERSION AX345904.1 GI:18493790
 KEYWORDS
 SOURCE
 ORGANISM
 synthetic construct.
 artificial sequences.
 REFERENCE
 1
 AUTHORS Olek,A., Plepenbrock,C. and Berlin,K.
 TITLE Diagnosis of diseases associated with the immune system
 JOURNAL Patent: WO 0200928-A 975 03-JAN-2002;
 Epigenomics AG (DE)
 FEATURES
 source location/Qualifiers
 1..8700
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="chemically treated genomic DNA (Homo sapiens)"
 BASE COUNT 2247 a 281 c 2233 g 3939 t
 ORIGIN
 Query Match 63.2%; Score 1141.4; DB 6; Length 8700;
 Best Local Similarity 80.0%; Pred. No. 1.7e-260;
 Matches 1454; Conservative 0; Mismatches 351; Indels 12; Gaps 9;
 QY 1 GGGGATTCGTTTGTCTTCAATTTTCAATGAATCAGAGATCTGTTGGGTG 60
 Db 3055 GGGGATTCGTTTGTCTTCAATTTTCAATGAATCAGAGATCTGTTGGGTG 3114
 QY 61 CAAAGCAGATCTAGAGAGGATGATACAGAGAAAG -AACAGCAAGCAGATTAAG 119
 Db 3115 TAAAGTATATTTAGAGAGGATGATTAAGAGAAAGAAATAGTAAAGCAGATTAAG 3174


```

||||| 5462 ATTCGTTTATTTCTTCAATTTTCCAAATTAATCAAAATCTATCTTAATATCAAC 5583
||||| 65 GCAGATAGTAGAGAGGTGATACAGAGAAAGAA-CAGCAAGCCAGATTATGGCAG 123
||||| 5582 GCAATACTAAAAAATAATATCAAAAAAACAACAAGCAGATTATACAG 5523
||||| 124 GTTCTCTGTAACAAAGTTGAGTGACACAGCTGAGCAGTGTGGAGAAAGCTCAT 183
||||| 5522 ATTTCTTAATAAATAATATATACCACTAAACCTATATATAAAAAATCAT 5463
||||| 184 AAGAAATAGAGGTGTGGCTGTGTCACCCGGGGCCCTCATGTCTGTCTTGG 243
||||| 5462 AAAAAATAGCATCTAAACCTGTGTCACCCGAAACCTCATTTCTTATCTTAA 5403
||||| 244 TCTCTTTTATTTAGAGTCCAAATTTATTTATTTAGTACAGAGGAGCAAT 303
||||| 5402 TCTCTTTTATTTATTAATAATCAATTTATTTATTTATTAATCAAAAAAAGCAAT 5343
||||| 304 GATCTTTCTTAAAGAGAGATATATATTAAGAGAGCTGTATAGATATGGG 363
||||| 5342 ATCTTTCTATTTAAAAAATAATATATATATATATATATATATATATATAT 5283
||||| 364 GAAGAGTGAGAGGGGAAAGGGAGAGAGAGAGAGAGAGAGAGAGAGAG 423
||||| 5282 AAAAAATATACAAAAAATAAAAAAAGAAAAAAGAAAAAAGAAAAAAG 5223
||||| 424 AAGAGAGACACTGGGCGAGAGATATAGAGAGAGAGAGAGAGAGAGAGAGAG 483
||||| 5222 AAAAAAAGCACTAAGCAAAAAATGATTAATAAAAAAATAATAATAAATTA 5163
||||| 484 CTTCACCCAGAGCTTCGTGCGTGGAGAGAGATGAGAGAGCTCTGATTTGCTAT 543
||||| 5162 CTTCACCCAGAGCTTCGT-ACATAAAAAAATAAAAAAATAAAAAAATAAAAA 5104
||||| 544 TCCAAATGCAAGAGGCTCTCAAGTCACTATCCACTCTTAAGAGAGAGGCTCAAT 603
||||| 5103 TCCAAATGCAAAAAATCTCTCAATCTCACTATCCACTCTTAAGAGAGGCTCAAT 5044
||||| 604 TCAACAGGCTCAGAAAGTCTCTCCGCGAGAGTCTCCAGCTCTCCAGTCCAC 663
||||| 5043 TCAACAGGCTCAGAAATATCTCTCCGCGAGAGTCTCCAGCTCTCCAGTCCAC 4984
||||| 664 ACAAACTCTTGGAAGAGTCTCTGGAATATTAATCTCAATCCAACTGAGACCA 723
||||| 4983 ACAAACTCTTGGAAGAGTCTCTGGAATATTAATCTCAATCCAACTGAGACCA 4924
||||| 724 GGTCTCTGTTGGTACCG-AGAGAGGGGTGCGCAGACAAATGAGAGAACTGAG 782
||||| 4923 AGTCTCTATTAATACCGAATAAAAAAATAATACGCAAACTAATAAAAACTG 4864
||||| 783 GCCAGAGAGGCGCAGAGAGTATACAGCAGCTCAGCGCAATTGCG-CCCCA 841
||||| 4863 ACCAAAAAAGCCGCAAAAAATTAACAGCAGCTCAAGCGCAATTAAGCGCCG 4804
||||| 842 ACTGAAAGAGTGTGATGAGAGATAGCTAGAACTCCCAAGCAGTGAAGATCA 901
||||| 4803 ACTGAAAGAGTGTGATGAGAGATAGCTAGAACTCCCAAGCAGTGAAGATCA 4744
||||| 902 ATGAGAGAGGCTCTCAAAAAATGACAAATGCAATGCAATTAAGAGCTATTTG 961
||||| 4743 ATGAGAGAGGCTCTCAAAAAATGACAAATGCAATGCAATTAAGAGCTATTTG 4684
||||| 962 ATTAGGTTTACAGAGTCCAGTATTTATTTATTTATTTATTTATTTATTTATTT 1021
||||| 4683 ATTAGGTTTACAGAGTCCAGTATTTATTTATTTATTTATTTATTTATTTATTT 4624
||||| 1022 TCCATTGAAAAATATCAGGAGATACCAAGAGAGAGAGAGAGAGAGAGAG 1081
||||| 4623 TCCATTGAAAAATATCAGGAGATACCAAGAGAGAGAGAGAGAGAGAGAGAG 4564
||||| 1082 ATATTGAAAAAGCTGCTTAAGA-GCATTCGCTTGCGGCTCCACTTAAATGGGA 1140
|||||

```

```

Db 4563 ATATTAAAAAACCTTACCTTAATAAACGATTTGGCTTACGATCCACTTTAATTA 4504
Oy 1141 CTTCGGGTGAGACACAGTGAAGAGTCTTCTGCTTGAAGAAAGCCT---GGAAG 1197
Db 4503 CTTCAGATATTAACACACAGTGAAGAGTCTTCTGCTTGAAGAAAGCCT---GGAAG 4444
Oy 1198 CGAAGGCCCCGGGCACTTTCAGATGCTATTTGTGGCCCTTGGGAGATTAACAGCC 1257
Db 4443 GAAAGCCCCGAGCAGCTTTCAGATGCTATTTGTAAACCCCTTAATAAATAAACA 4384
Oy 1258 AGGGGTGAATTAACCCCGAGTCCCTGCTGCTGAGACCAATGTAAGTCAG 1317
Db 4383 AAGGATATTAATTAACCCCGAGTCCCTGCTGCTGAGACCAATGTAAGTCAG 4324
Oy 1318 AATGTCCTCAAGCTTCCCTGCTGCTGAGATTAATTAAGAGTCCAGAGTACAC 1377
Db 4323 AATGTCCTCAAGCTTCCCTGCTGCTGAGATTAATTAAGAGTCCAGAGTACAC 4264
Oy 1378 AGAATGCGAGCGCTGTATGCCAGTTCCCGACACCGGCTGCGCAGAGAGACCTCA 1437
Db 4263 AGAATGCGAGCGCTGTATGCCAGTTCCCGACACCGGCTGCGCAGAGAGACCTCA 4204
Oy 1438 CCCCCGAGCGGGAAGGGGTAAGCGCGCGGGTCAAGAGATGCGGGGTCTAGTTGGC 1497
Db 4203 CCCCCGAGCGGGAAGGGGTAAGCGCGCGGGTCAAGAGATGCGGGGTCTAGTTGGC 4144
Oy 1498 CAGAGTGAAGTGGGTGACCGCGGGGTGCTGAGGTGCTGAGTGGCGGGGTGCGGG 1557
Db 4143 CAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4084
Oy 1558 ACACTTGTCTTGTATGACAGCAGAGTGA--CGGCTCGGGCTGAGCAGAGAGCA 1615
Db 4083 ACACTTGTCTTGTATGACAGCAGAGTGA--CGGCTCGGGCTGAGCAGAGAGCA 4024
Oy 1616 GGTGCGGGGCGAGTCTGAGTGCAGCGCCCGGAGCTGCGGGCTGCTGAGGCGCA 1675
Db 4023 AATACCGCAAGATTCGACTACAGCGCCCGAAGTCAAGCCGCACTACTAATAACGA 3964
Oy 1676 AGCAGC-GGCGCGCGCAGCGCTGCTGCGCGAGCGCCCGCTCCCAACCCCACTCTG 1734
Db 3963 AAGCAGCAGCGCGCGCAGCGCTGCTGCGCGAGCGCCCGCTCCCAACCCCACTCTG 3904
Oy 1735 GCGCGC-CGTTCCGGGGGTGCTGCGCGAGCGCGCGCTTCTATATAGCGGGCGCGC 1793
Db 3903 AAGCGCGCGTTCGGAACGATCTTAACACCCCGCACTTCTATATAGCAGCGCGCG 3844
Oy 1794 CCGGCGCGCGCA 1805
Db 3843 CCGAACCGCGCA 3832

RESULT 7
AF305712 13720 bp DNA linear MAM 05-NOV-2000
LOCUS Bos laurus insulin-like growth factor binding protein-3 (IGFBP3)
DEFINITION Bos laurus insulin-like growth factor binding protein-3 (IGFBP3)
ACCESSION AF305712
VERSION AF305712.1 GI:11095302
KEYWORDS
SOURCE
ORGANISM
Bos laurus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
Bovidae; Bovinae; Bos.
1 (bases 1 to 13720)
Erondou,N.E., Toland,B., Boes,M., Dake,B., Moser,D.R. and Bar,R.S.
Bovine insulin-like growth factor binding protein-3: organization
of the chromosomal gene and functional analysis of its promoter
Endocrinology 138 (7), 2856-2862 (1997)
JOURNAL MEDLINE
PUBMED 97345736
9202228
2 (bases 1 to 13720)
Erondou,N.E., Toland,B., Boes,M., Dake,B. and Bar,R.S.
REFERENCE
AUTHORS

```

TITLE Direct Submission
JOURNAL Submitted (14-SEP-2000) Diabetes and Endocrinology Research Center

Department of Internal Medicine, The University of Iowa and
Veteran's Administration Medical Center, 601 Highway 6 West, Iowa
City, IA 52246, USA

exon	4336..4874	/product=insulin like growth factor binding protein-3
	/gene="IGFBP3"	
	/number=1	
CDS	join(4472..4874,7829..8055,8550..8649,9699..9824)	

polyA_signal	12769.	.12774		
	/gene="IGFBP3"			
BASE COUNT	3369 a	3479 c	3393 g	3479 t
ORIGIN				

	Query Match	10.5%	Score 190;	DB 4;	Length 13720;
	Best Local Similarity	57.3%	Pred. No. 3e-34;		
	Matches	800;	Conservative	0;	Mismatches 465; Indels 131; Gaps 20;
QY	507	GAGGCGAATGAGAGAGCTCTGATTTGCTATTACTCTCAAACTGCAGAAAGGCTCTTC	566		
Db	2972	GAGGGAAGCGAAGGAGAGCTCATTTGATGTTTCACTGCAAACTGCGCAAAAGGCTCTTC	3031		
QY	567	AAGTACCTATTCACCTCTCTA--AGGCAAGGCTCAATTGCAACGGTTCAGGAAAGTC	624		
Db	3032	TAGATGCTATCTCTCTCTAGCCGGGAGATATCCGTTTCCGACAGCTCCAGGAGACTT	3091		
QY	635	TCTCCCGCGAGGTCACGCGTTTCCACATCCACCCGCCAACAACTCTT--GGAAAG	681		
Db	3092	CGCTCCGCTGCTATGCGGGGTGACACTGTCATCCACAGGCTCTTTTAAAGAACT	3151		
QY	682	TGCCCTGAAAAATTAATATCTCAATCAAACTCTGGACCAACGCGCTCTGTGGTAC	741		
Db	3152	TGAAGGTATTGATTAATTTCCCAATTCATCTGGACCACTGGCGCTCTTGTGGTAC	3211		
QY	742	CGAAGGAGGGGTGGCGAGACAAACTGAGAAACTCGAGTCCGAGAGAGGCCACAGG	801		

31151 32640: contig of 1490 bp in length
32641 32740: gap of unknown length
32741 34019: contig of 1279 bp in length
34020 34119: gap of unknown length
34120 36105: contig of 1986 bp in length
36106 36205: gap of unknown length
36206 37278: contig of 1073 bp in length
37279 37379: gap of unknown length
37380 39416: contig of 2038 bp in length
39417 41021: contig of 1505 bp in length
41022 41121: gap of unknown length
41122 42627: contig of 1506 bp in length
42628 42727: gap of unknown length
42728 44369: contig of 1642 bp in length
44370 44469: gap of unknown length
44470 45561: contig of 1092 bp in length
45562 45661: gap of unknown length
45662 46976: contig of 1315 bp in length
46977 47076: gap of unknown length
47077 48433: contig of 1267 bp in length
48434 49674: contig of 1231 bp in length
49675 49774: gap of unknown length
49775 51642: contig of 1868 bp in length
51643 51742: gap of unknown length
51743 53097: contig of 1355 bp in length
53098 53197: gap of unknown length
53198 55117: contig of 1920 bp in length
55118 55217: gap of unknown length
55218 57573: contig of 2356 bp in length
57574 57673: gap of unknown length
57674 59663: contig of 1890 bp in length
59664 61822: contig of 2159 bp in length
61823 61922: gap of unknown length
61923 63079: contig of 1157 bp in length
63080 63179: gap of unknown length
63180 64615: contig of 1436 bp in length
64616 64715: gap of unknown length
64716 67307: contig of 2592 bp in length
67308 67407: gap of unknown length
67408 68773: contig of 1366 bp in length
68774 68873: gap of unknown length
68874 70712: contig of 1839 bp in length
70713 70812: gap of unknown length
70813 73004: contig of 2192 bp in length
73005 73104: gap of unknown length
73105 74453: contig of 1349 bp in length
74454 74553: gap of unknown length
74554 76324: contig of 1771 bp in length
76325 76424: gap of unknown length
76425 77737: contig of 1313 bp in length
77738 77837: gap of unknown length
77838 79745: contig of 1908 bp in length
79746 79845: gap of unknown length
79846 81448: contig of 1603 bp in length
81449 81548: gap of unknown length
81549 83338: contig of 1790 bp in length
83339 83438: gap of unknown length
83439 84877: contig of 1433 bp in length
84878 84977: gap of unknown length
84978 87020: contig of 2043 bp in length

Query Match 4.6%; Score 82.6; DB 2; Length 107161;
Best Local Similarity 60.0%; Pred. No. 1.4e-08;
Matches 219; Conservative 0; Mismatches 124; Indels 22; Gaps 4;

QY 1445 AGCGAAGGAGGTAAGGCGGCGGCTCAAGAGATCGGGGCTGAGTTGCCAGGAGT 1504
DB 67899 AGAGGAGGAGGCGTAAGGCGGCGGAGCCCAATTTGGCTAAAGTCTTTGGGGAGAGAG 67840
QY 1505 GACTGGGGTGACCGGGGCTGCTGAGTGCCTGAGTCCGCGGGGTGGCCGGCACACCT- 1563

DB 67839 GAACAGAGCCCCGATGATGCTGGGACTCAGAGACAGCAAGTAGCCTGGAGCCCTG 67780
QY 1564 TGTCTTCTTGTAGAGACAAAGTG--ACGGGCTCCGGCGTGGAGACAGAGAGAGTGGC 1621
DB 67779 CCGTTACCGGGGAGCAAGAAAGTGGCCCTAGTCTCTGGAGTCCAGAGAGAGAGTGGC 67720
QY 1622 CGGGCGAGTCTGAGCTGACAGAGCCCGGAGCTGGCCCGGCTGCTCAGAGCGAGAGCAG 1681
DB 67719 CAGTGAAGACAGCGGAGCGGCGGCGCCGCTGCTGATCCGCTTCGCGCG--GAGGCGGG 67663
QY 1682 GCGCCCGAGCCGCTGCTGGCGGAGCCCGGCTGCTCCAGCCCACTCTCTGGCGCC 1741
DB 67662 AGCCCGGAGCGGAGCTGCTGGCGGAGCCCGGCTGCTCCAGCCCGCGC-----C 67619
QY 1742 GTTCGGGAGGCTGCTGCTGGCGGAGCCCGGCTGCTATATACGGCGGCGGCGCGG 1801
DB 67618 GTTCGGGAGGCTGCTGCTGGCGGAGCCCGGCTTATATATGAGCTGCGGCGGAGCTG 67559
QY 1802 CCCAG 1806
DB 67558 CCCAG 67554

RESULT 9
LOCUS 166494/c 7218 bp DNA linear PAT 28-DEC-1997
DEFINITION Sequence 14 from patent US 5670367.
ACCESSION 166494
VERSION 166494.1 GI:2724471
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7218)
AUTHORS Dörner, F., Scheiflinger, F. and Falkner, F. Gunter.
TITLE Recombinant fowlpox virus
JOURNAL Patent: US 5670367-A 14 23-SEP-1997;
FEATURES
source 1. 7218
location/Qualifiers
BASE COUNT 1944 a 1491 c 1486 g 1929 t 368 others
ORIGIN

Query Match 4.4%; Score 78.6; DB 6; Length 7218;
Best Local Similarity 11.2%; Pred. No. 8.3e-08;
Matches 30; Conservative 159; Mismatches 78; Indels 0; Gaps 0;

QY 256 TGTAGAGCTCAATTATTTATTTAGTACAGAGGAGCAAAATGATCTTCCATT 315
DB 1496 TGTAGGATCATCTGATATTTACTATCTATGACAGTAGTTAAAGAGATAGAAATTTGGT 1437
QY 316 CTAAAGAGAGATATATGATATAAAGAGCTGATAGATATGGGGAAGAGTGAGC 375
DB 1436 ACRR 1377
QY 376 AGGGGGAAGAGGAG 435
DB 1376 RRR 1317
QY 436 TGGCGAGAGATGATTTAGAGAGACAGAAATGATGATGAAGATTAACTGACCCAGG 495
DB 1316 RRR 1257
QY 496 CTTCGCTGCGAGAGGAGATGAGAGG 522
DB 1256 RRR 1230

RESULT 10
LOCUS AC078884 207420 bp DNA linear HTG 15-MAY-2002
DEFINITION Mus musculus chromosome 17 clone RP23-390F1 strain C57BL6/J,

[illegible]

Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified, using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RN/RepeatMasker.html>
Genome Center
Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L1/376
Center clone name: 100_D_4

* NOTE: This record contains 85 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.

* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
725 824: gap of 100 bp in length
825 1513: contig of 689 bp in length
1514 1613: gap of 100 bp
1614 2330: contig of 717 bp in length
2331 2430: gap of 100 bp
2431 3129: contig of 699 bp in length
3130 3229: gap of 100 bp
3230 3928: contig of 699 bp in length
3929 4028: gap of 100 bp
4029 4754: contig of 726 bp in length
4755 4854: gap of 100 bp
4855 5567: contig of 713 bp in length
5568 5667: gap of 100 bp
5668 6396: contig of 729 bp in length
6397 6496: gap of 100 bp
6497 7224: contig of 728 bp in length
7225 7324: gap of 100 bp
7325 8038: contig of 714 bp in length
8039 8138: gap of 100 bp
8139 8846: contig of 708 bp in length
8847 8946: gap of 100 bp
8947 9683: contig of 737 bp in length
9684 9783: gap of 100 bp
9784 10482: contig of 699 bp in length
10483 10582: gap of 100 bp
10583 11306: contig of 724 bp in length
11307 11406: gap of 100 bp
11407 12125: contig of 719 bp in length
12126 12225: gap of 100 bp
12226 12936: contig of 711 bp in length
12937 13036: gap of 100 bp
13037 13744: contig of 708 bp in length
13745 13844: gap of 100 bp
13845 14553: contig of 709 bp in length
14554 14653: gap of 100 bp
14654 15362: contig of 709 bp in length
15363 15462: gap of 100 bp
15463 16183: contig of 721 bp in length
16184 16283: gap of 100 bp
16284 17005: contig of 722 bp in length
17006 17105: gap of 100 bp
17106 17824: contig of 719 bp in length
17825 17924: gap of 100 bp
17925 18628: contig of 704 bp in length
18629 18728: gap of 100 bp
18729 19440: contig of 712 bp in length
19441 19540: gap of 100 bp
19541 20261: contig of 721 bp in length
20262 20361: gap of 100 bp
20362 21061: contig of 700 bp in length
21062 21161: gap of 100 bp
21162 21876: contig of 715 bp in length
21877 21976: gap of 100 bp
21977 22677: contig of 701 bp in length
22678 22777: gap of 100 bp
22778 23507: contig of 730 bp in length
23508 23607: gap of 100 bp
23608 24317: contig of 710 bp in length
24318 24417: gap of 100 bp
24418 25128: contig of 711 bp in length
25129 25228: gap of 100 bp
25229 25944: contig of 716 bp in length
25945 26044: gap of 100 bp
26045 26772: contig of 728 bp in length
26773 26872: gap of 100 bp
26873 27576: contig of 704 bp in length
27577 27676: gap of 100 bp
27677 28377: contig of 701 bp in length

28378 28477: gap of 100 bp
28478 29181: contig of 704 bp in length
29182 29281: gap of 100 bp
29282 30001: contig of 720 bp in length
30002 30101: gap of 100 bp
30102 30822: contig of 721 bp in length
30823 30922: gap of 100 bp
30923 31646: contig of 724 bp in length
31647 31746: gap of 100 bp
31747 32463: contig of 717 bp in length
32464 32563: gap of 100 bp
32564 33281: contig of 718 bp in length
33282 33381: gap of 100 bp
33382 34085: contig of 704 bp in length
34086 34185: gap of 100 bp
34186 34910: contig of 725 bp in length
34911 35010: gap of 100 bp
35011 35729: contig of 719 bp in length
35730 35829: gap of 100 bp
35830 36547: contig of 718 bp in length
36548 36647: gap of 100 bp
36648 37344: contig of 697 bp in length
37345 37444: gap of 100 bp
37445 38149: contig of 705 bp in length
38150 38249: gap of 100 bp
38250 38956: contig of 707 bp in length
38957 39056: gap of 100 bp
39057 39766: contig of 710 bp in length
39767 39866: gap of 100 bp
39867 40563: contig of 697 bp in length
40564 40663: gap of 100 bp
40664 41387: contig of 724 bp in length
41388 41487: gap of 100 bp
41488 42206: contig of 719 bp in length
42207 42306: gap of 100 bp
42307 42982: contig of 676 bp in length
42983 43082: gap of 100 bp
43083 43787: contig of 705 bp in length
43788 43887: gap of 100 bp
43888 44597: contig of 710 bp in length
44598 44697: gap of 100 bp
44698 45423: contig of 726 bp in length
45424 45523: gap of 100 bp
45524 46238: contig of 715 bp in length
46239 46338: gap of 100 bp
46339 47056: contig of 718 bp in length
47057 47156: gap of 100 bp
47157 47868: contig of 712 bp in length
47869 47968: gap of 100 bp
47969 48679: contig of 711 bp in length
48680 48779: gap of 100 bp
48780 49483: contig of 704 bp in length
49484 49583: gap of 100 bp
49584 50268: contig of 685 bp in length
50269 50368: gap of 100 bp
50369 51092: contig of 724 bp in length
51093 51192: gap of 100 bp
51193 51909: contig of 717 bp in length
51910 52009: gap of 100 bp
52010 52713: contig of 704 bp in length
52714 52813: gap of 100 bp
52814 53502: contig of 689 bp in length
53503 53602: gap of 100 bp
53603 54330: contig of 728 bp in length
54331 54430: gap of 100 bp
54431 55141: contig of 711 bp in length
55142 55241: gap of 100 bp
55242 55961: contig of 720 bp in length
55962 56061: gap of 100 bp

Query Match 4.1% Score 74.2; DB 2; Length 68866;
Best Local Similarity 41.4%; Pred. No. 1.3e-06;
Matches 164; Conservative 0; Mismatches 229; Indels 3; Gaps 1;

Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
 Delaney, K.R., Delgado, O., Deen, A.L., Ding, Y., Dinh, H.H.,
 Douthett, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
 Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
 Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B.,
 Homel, F., Howard, S., Huber, J., Huliy, S., Hume, J., Jackson, L.E.,
 Jacobson, B., Jia, Y., Johnson, R., Jolyet, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, U., King, U., Kovah, J., Kovar, C.,
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
 Li, J., Li, Z., Licharge, O., Lieu, C., Liu, J., Liu, W., Lousseng, H.,
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
 Massey, E., Mayhew, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, K., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, E., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G.,
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
 Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,
 Scherer, S., Scott, G., Shen, H., Shooshbari, N., Slisdon, I.,
 Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wlezyk, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G., and Gibbs, R.

Unpublished
 Direct Submission
 2 (bases 1 to 187521)
 Worley, K.C.
 Submitted (29-JUN-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 187521)
 Worley, K.C.
 Direct Submission
 Submitted (24-JUN-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jun 30, 2002 this sequence version replaced gi:21629151.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

----- Project Information
 Center project name: GZEE
 Center clone name: CH230-204D14

----- Summary Statistics
 Sequencing vector: Plasmid
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap, version 0.990329
 Consensus quality: 131060 bases at least Q40
 Consensus quality: 138483 bases at least Q30
 Consensus quality: 143840 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 73 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will

* be preserved.
 1
 1213
 1313
 2942
 2942
 3042
 4407
 4507
 4507
 6177
 6277
 7654
 7754
 9023
 9123
 10550
 10650
 11655
 11765
 11765
 12783
 12883
 13946
 14046
 15825
 15925
 17209
 17309
 18987
 19087
 20250
 20350
 22358
 22458
 23672
 23772
 25527
 25627
 26803
 26804
 26904
 28445
 28445
 30094
 30194
 32190
 32290
 33290
 33405
 33406
 34911
 35011
 35011
 36646
 36746
 38285
 38286
 40885
 40885
 40985
 42862
 43062
 44758
 44858
 46199
 46299
 46399
 48625
 48725
 50474
 50574
 52657
 52757
 54755
 54855
 56445
 56446
 58340
 58441
 59626

1212: contig of 1212 bp in length
 1312: gap of unknown length
 2941: contig of 1629 bp in length
 2941: gap of unknown length
 4406: contig of 1365 bp in length
 4506: gap of unknown length
 4506: contig of 1670 bp in length
 6176: gap of unknown length
 6276: contig of 1377 bp in length
 7653: gap of unknown length
 7753: gap of unknown length
 9022: contig of 1265 bp in length
 9122: gap of unknown length
 10549: contig of 1421 bp in length
 10649: gap of unknown length
 11664: contig of 1015 bp in length
 11764: gap of unknown length
 12782: contig of 1018 bp in length
 12882: gap of unknown length
 13945: contig of 1063 bp in length
 14045: gap of unknown length
 15824: contig of 1779 bp in length
 15924: gap of unknown length
 17208: contig of 1284 bp in length
 17308: gap of unknown length
 18986: contig of 1678 bp in length
 19086: gap of unknown length
 20249: contig of 1163 bp in length
 20349: gap of unknown length
 22357: contig of 2008 bp in length
 22457: gap of unknown length
 23671: contig of 1214 bp in length
 23771: gap of unknown length
 25527: contig of 1756 bp in length
 25627: gap of unknown length
 26803: contig of 1176 bp in length
 26804: gap of unknown length
 26904: contig of 1542 bp in length
 28445: gap of unknown length
 28445: gap of unknown length
 30093: contig of 1548 bp in length
 30193: gap of unknown length
 32189: contig of 1996 bp in length
 32289: gap of unknown length
 33290: contig of 1016 bp in length
 33405: gap of unknown length
 33406: gap of unknown length
 34910: contig of 1505 bp in length
 35010: gap of unknown length
 36645: contig of 1635 bp in length
 36745: gap of unknown length
 38285: contig of 1540 bp in length
 38285: gap of unknown length
 40884: contig of 2499 bp in length
 40984: gap of unknown length
 42861: contig of 1977 bp in length
 43061: gap of unknown length
 44757: contig of 1696 bp in length
 44857: gap of unknown length
 46198: contig of 1341 bp in length
 46298: gap of unknown length
 48624: contig of 2326 bp in length
 48724: gap of unknown length
 50473: contig of 1749 bp in length
 50573: gap of unknown length
 52656: contig of 2083 bp in length
 52756: gap of unknown length
 54754: contig of 1998 bp in length
 54854: gap of unknown length
 56445: contig of 1591 bp in length
 56445: gap of unknown length
 58340: contig of 1795 bp in length
 58440: gap of unknown length
 59625: contig of 1165 bp in length
 59725: gap of unknown length

```
* 59726 62021: contig of 2296 bp in length
* 62022 62121: gap of unknown length
* 62122 63651: contig of 1530 bp in length
* 63652 66236: gap of unknown length
* 66237 66336: contig of 2485 bp in length
* 66337 66540: gap of unknown length
* 66541 68640: contig of 2204 bp in length
* 68641 70800: gap of unknown length
* 70801 73422: contig of 2160 bp in length
* 73423 73522: gap of unknown length
* 73523 76025: contig of 2502 bp in length
* 76025 76125: gap of unknown length
* 76125 77513: contig of 1389 bp in length
* 77514 79925: gap of unknown length
* 79926 80025: contig of 2312 bp in length
* 80026 82185: gap of unknown length
* 82185 82284: contig of 2159 bp in length
* 82285 83574: gap of unknown length
* 83575 85573: contig of 1290 bp in length
* 85574 85673: gap of unknown length
* 85674 88158: contig of 1899 bp in length
* 88159 88258: gap of unknown length
* 88259 91411: contig of 2485 bp in length
* 91412 91511: gap of unknown length
* 91512 94235: contig of 5153 bp in length
* 94236 94335: gap of unknown length
* 94336 96689: contig of 2724 bp in length
* 96690 96789: gap of unknown length
* 96790 100020: contig of 2354 bp in length
* 100021 100120: gap of unknown length
* 100121 102772: contig of 3231 bp in length
* 102772 102772: contig of 2652 bp in length

Query Match      3.8%; Score 68.2; DB 2; Length 187521;
Best Local Similarity 47.6%; Pred. No. 4.2e-05;
Matches 166; Conservative 0; Mismatches 183; Indels 0; Gaps 0;
```

```
AL591129/c 213985 bp DNA linear ROD 29-JUN-2002
LOCUS Mouse DNA sequence from clone Rp23-147P4 on chromosome 11, complete
DEFINITION sequence.
ACCESSION AL591129
VERSION AL591129.24 GI:21665918
KEYWORDS HTG.
```

SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Fukuyama; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 213985)
AUTHORS Clark, G.
TITLE Direct Submission
JOURNAL Submitted (29-JUN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonequest@sanger.ac.uk
On Jul 2, 2002 this sequence version replaced gi:2165332.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquerry@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep Rp23-147P4 is from the Rp23-Mouse PAC Library

constructed by the group of Pieter de Jong.

For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6.

FEATURES
source location/Qualifiers
1..213985
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/clone="Rp23-147P4"
/clone_1lb="Rp23-23"

BASE COUNT 62282 a 47467 c 48155 g 56081 t
ORIGIN

Query Match 3.8%; Score 68.2; DB 10; Length 213985;
Best Local Similarity 56.4%; Pred. No. 4.3e-05;
Matches 127; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

```
QY 256 TGTAGAGGCTCAATTTTATTATTAGTACAAAGGAGAGAAATGATTTCCATT 315
|||||
Db 165045 TGTAGAGGCTGTGATCTTGAGTTTGTCAGTCCTCAATCTGAGACCTGT 164986
QY 316 CTAAGAGAGATATATATTAAGAGAGCTGTATAGATATGGGAGAGGTGAC 375
|||||
Db 164985 CTCAAAAAGAGAGATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 164926
QY 376 AGGGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 435
|||||
Db 164925 AGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 164866
QY 436 TGGGCGAGAGATGATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 480
|||||
Db 164865 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 164821
```

```
RESULT 15
AL806529/c 214511 bp DNA linear HTG 17-AUG-2002
LOCUS AL806529/c
```


GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 23, 2003, 06:54:28 ; Search time 408 Seconds
(without alignments)
9968.404 Million cell updates/sec

Title: US-09-865-879-1

Perfect score: 1806
Sequence: 1 gggagctcgttttcttccct.....ggcgcccgagcgccag 1806

Scoring table: IDENTITY_MUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: N_Geneseq_101002:*
2: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
3: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
4: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
5: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
6: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
7: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
8: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
9: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
10: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
11: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
12: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
13: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
14: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
15: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
16: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
17: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
18: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
19: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
20: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
21: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
22: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
23: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
24: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
25: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the total score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1806	100.0	1806	24	AAD24912
2	1795	99.4	10884	24	ABN95848
3	1141.4	63.2	8700	24	ABL33002
4	1021.2	56.5	8700	24	ABL33003
5	117	6.5	1007	24	ABO45156
6	117	6.5	1007	24	ABO45157
7	84.4	4.7	1007	24	ABO45154
8	84.4	4.7	1007	24	ABO45155
9	60.8	3.4	6027	24	ABN80173

10	59	3.3	395	22	AAF64375	Novel human polynu
11	59	3.3	16200	24	AAD36438	Mouse L66 genomic
12	58.4	3.2	170	21	ABN81171	Shrimp polynucleot
13	57.8	3.2	502	21	ABN81079	Shrimp polynucleot
14	57.4	3.2	396	22	AAF64796	Novel human polynu
15	57	3.2	420	22	AAF67104	Novel human polynu
16	56.6	3.1	307	24	ABL86438	Novel human polynu
17	56.6	3.1	5191	21	AAZ98617	Human ovarian can
18	56.6	3.1	10301	22	AAK84949	RPPI-WSA genomic n
19	56.2	3.1	329	17	AAI12183	Human immune/haema
20	56.2	3.1	484	22	AAF64952	Partial p3G4-5-CDK
21	56.2	3.1	421	22	AAF64755	Novel human polynu
22	56.2	3.1	36301	20	AAZ23892	Novel human polynu
23	56.2	3.1	38886	20	AAZ23897	Murine LOB0 genom
24	56.2	3.1	611590	21	AAF22303	Murine LOB0 homolo
25	56	3.1	395	22	AAF66094	Arabidopsis thailia
26	56	3.1	611590	21	AAF22303	Arabidopsis thailia
27	55.6	3.1	298	24	ABL86000	Novel human polynu
28	55.6	3.1	397	22	AAF64573	Human ovarian can
29	55.6	3.1	863	20	AAZ17377	Novel human polynu
30	55.6	3.1	3586	24	AAZ63366	Human gene express
31	55.4	3.1	398	22	AAF64279	Chemically pretrea
32	55.4	3.1	400	22	AAF64260	Novel human polynu
33	55.4	3.1	422	22	AAF66607	Novel human polynu
34	55.4	3.1	424	22	AAF67052	Novel human polynu
35	55.2	3.1	390	22	AAF64933	Novel human polynu
36	55.2	3.1	399	22	AAF66668	Novel human polynu
37	55.2	3.1	405	22	AAF64515	Novel human polynu
38	55.2	3.1	4752	15	AAO58032	Tyrosinase promote
39	55	3.0	5306	24	ABN32510	Human immune syste
40	54.8	3.0	190	16	AAI45061	Ts2 gene microsat
41	54.6	3.0	300	20	AAI13219	Human gene express
42	54.6	3.0	382	22	AAF65098	Novel human polynu
43	54.6	3.0	700	11	AAO05263	Sequence of neuron
44	54.6	3.0	700	15	AAO57505	Rat GAP-43 promote
45	54.6	3.0	700	15	AAO57506	Rat GAP-43 promote

ALIGNMENTS

RESULT 1	
AAD24912	
ID	AAD24912 standard; DNA; 1806 BP.
XX	
AC	AAD24912;
XX	
DT	12-MAR-2002 (first entry)
XX	
DE	Human insulin-like growth factor binding protein-3 promoter DNA.
XX	
KW	Human: growth inhibitory gene; retinoid; retinoid acid response element;
KW	RARE site; therapy; promyelocytic leukaemia; cancer chemoprevention;
KW	cytostatic; insulin-like growth factor binding protein; IGFBP-3 promoter;
KW	ds.
OS	
XX	
XX	Homo sapiens.
XX	
PN	WO200192578-A2.
XX	
PD	06-DEC-2001.
XX	
PF	25-MAY-2001; 2001WO-US17161.
XX	
PR	26-MAY-2000; 2000US-207535P.
XX	
PA	(UNIT) UNITV ILLINOIS FOUNO.
XX	
PI	Roninson IB, Dokmanovic M, Chang B;
XX	
DR	WPI; 2002-075474/10.
XX	
PT	Expression construct encoding cellular genes, under control of a

PT promoter regulated by retinoids and cells comprising the construct for
PT identifying compounds that induce expression of the genes useful in
PT treating cancer -

XX Claim 4: Page 42-43; 64pp; English.

XX The patent discloses growth inhibitory genes induced by retinoids. The
CC invention also relates to recombinant expression constructs that express
CC a reporter gene under the transcriptional control of a promoter for a
CC gene which is expressed by retinoid induction. The promoter does not
CC contain a retinoic acid response elements (RARE) site. The invention
CC further relates to reagents and methods for identifying compounds other
CC than retinoids that modulate the expression of cellular genes. These
CC compounds are useful for treating cancers such as promyelocytic leukemia
CC and cancer chemoprevention. The present sequence is human insulin-like
CC growth factor binding protein (IGFBP-3) promoter DNA used in the
CC invention.

XX Sequence 1806 BP; 467 A; 437 C; 517 G; 385 T; 0 other;

Query Match 100.0%; Score 1806; DB 24; Length 1806;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1806; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGATTCGTTTGTTCCTCAATTTCCAAATGAATCAGAGATCCTGTTGGGTGT 60
DB 1 GGGGATTCGTTTGTTCCTCAATTTCCAAATGAATCAGAGATCCTGTTGGGTGT 60
QY 61 CAACGAGATCTAGAAAGAGGTGATCAAGAGAAAGAGAGAGAGAGAGATATGGC 120
DB 61 CAACGAGATCTAGAAAGAGGTGATCAAGAGAAAGAGAGAGAGAGAGATATGGC 120
QY 121 ACGGTTTCCTGTAAGAGAGGTGATCAAGAGAGAGAGAGAGAGAGAGAGT 180
DB 121 ACGGTTTCCTGTAAGAGAGGTGATCAAGAGAGAGAGAGAGAGAGAGAGT 180
QY 181 CATAGAGAAATGACGCTGCTGCGTCCACCGGCGCTCCATTTGTTGCTT 240
DB 181 CATAGAGAAATGACGCTGCTGCGTCCACCGGCGCTCCATTTGTTGCTT 240
QY 241 TGGTCTCTTTTATTTAGAGGTCATATTTATTTATTTAGTACAGAGAGAGAA 300
DB 241 TGGTCTCTTTTATTTAGAGGTCATATTTATTTATTTAGTACAGAGAGAGAA 300
QY 301 ATTTGATCTTTCATTTCAATTAAGAGAGATATATATTAAGAGAGAGATATG 360
DB 301 ATTTGATCTTTCATTTCAATTAAGAGAGATATATATTAAGAGAGAGATATG 360
QY 361 GGGGAGAGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 361 GGGGAGAGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 GACAGAGAGAGACATGCGGCGAGAGATCATTTAGAGAGAGAGAGAGAGAT 480
DB 421 GACAGAGAGAGACATGCGGCGAGAGATCATTTAGAGAGAGAGAGAGAT 480
QY 481 TAACCTTACCCCAAGGCTGCTGCTGAGAGGAGATGAGAGATCTGATTTGCTAT 540
DB 481 TAACCTTACCCCAAGGCTGCTGCTGAGAGGAGATGAGAGATCTGATTTGCTAT 540
QY 541 TACTCCAAACTGCAAGAGGCTCTTCAAGTACCTTATCCACTCTAAGCAAGCTCA 600
DB 541 TACTCCAAACTGCAAGAGGCTCTTCAAGTACCTTATCCACTCTAAGCAAGCTCA 600
QY 601 ATTTCAGAGCGTTTCAAGAAAGTCTCCCGGAGAGTTCACCGGTTCCACTCCAC 660
DB 601 ATTTCAGAGCGTTTCAAGAAAGTCTCCCGGAGAGTTCACCGGTTCCACTCCAC 660
QY 661 CCCACAACTCTTTGAAAAAGTCTTGAATAATTTATCTCAATCAATCTGAGACA 720
DB 661 CCCACAACTCTTTGAAAAAGTCTTGAATAATTTATCTCAATCAATCTGAGACA 720
QY 721 CGAGCGTCTCTGTTGGTACCGAGAGAGGAGGTGCGCAGACAAACTGAAGAACTCGA 780

DB 721 CCACGCTCCTCTGTTGGTACCGAGAGAGGAGGTGCGCAGACAAACTGAAGAACTCGA 780
QY 781 GTGCCAGAGAGAGGCGGACAGAGATTAAGGAGCTCAGGCGGCAATTTGGCCCGCACT 840
DB 781 GTGCCAGAGAGAGGCGGACAGAGATTAAGGAGCTCAGGCGGCAATTTGGCCCGCACT 840
QY 841 TACTGAAAGTGTATTAATTGACAGAGATTAAGTGAATCCCAAGCATGAGAAATCACT 900
DB 841 TACTGAAAGTGTATTAATTGACAGAGATTAAGTGAATCCCAAGCATGAGAAATCACT 900
QY 901 AATTCGAGAGTGCCTTCAAAAAATGACATGAATAATTCCTATTAAGAGACTATTTGGTT 960
DB 901 AATTCGAGAGTGCCTTCAAAAAATGACATGAATAATTCCTATTAAGAGACTATTTGGTT 960
QY 961 AATTAGCTTCAGAGAGGCGGAGTATTTGCTTTATTTATTTGCTGCTGGGTGTA 1020
DB 961 AATTAGCTTCAGAGAGGCGGAGTATTTGCTTTATTTATTTGCTGCTGGGTGTA 1020
QY 1021 CTCCATTTGAAAAATATATCAGGAGAGATATCCCAAGAGAGAGAGATTTGATTTTA 1080
DB 1021 CTCCATTTGAAAAATATATCAGGAGAGATATCCCAAGAGAGAGAGATTTGATTTTA 1080
QY 1081 AATTTTGAAGAGGCTGCTTAAAGAGATTCGTTGCGGCTCACCTTTAATTTGGGA 1140
DB 1081 AATTTTGAAGAGGCTGCTTAAAGAGATTCGTTGCGGCTCACCTTTAATTTGGGA 1140
QY 1141 CTTGCGGTTAGCAACGAGAGAGTCTTCTGTTGAGAAATTAAGCTTGAAGAGAGCA 1200
DB 1141 CTTGCGGTTAGCAACGAGAGAGTCTTCTGTTGAGAAATTAAGCTTGAAGAGAGCA 1200
QY 1201 AGGCGCGGCGCATCTTCAAGATCGTATTTGAGGCGGCTGAGGATATTAAGAGAGCA 1260
DB 1201 AGGCGCGGCGCATCTTCAAGATCGTATTTGAGGCGGCTGAGGATATTAAGAGAGCA 1260
QY 1261 GGGTGTAAATTAACCCCGAGTGCCTTGGCTCCGAGACCCAAATGTAAGTCAAGAAAT 1320
DB 1261 GGGTGTAAATTAACCCCGAGTGCCTTGGCTCCGAGACCCAAATGTAAGTCAAGAAAT 1320
QY 1321 GTCCCAAGACTTGGGCTGCGCAAGAGAAATTAATTAAGAAAGTCCAGAGAGTACACAG 1380
DB 1321 GTCCCAAGACTTGGGCTGCGCAAGAGAAATTAATTAAGAAAGTCCAGAGAGTACACAG 1380
QY 1381 AATGCGGAGCGCTGTAAGCAGTTTCCCGACACCGGCTCGCGGAGAGAGACTCACCC 1440
DB 1381 AATGCGGAGCGCTGTAAGCAGTTTCCCGACACCGGCTCGCGGAGAGAGACTCACCC 1440
QY 1441 CGAGAGCGGAAAGGCTTAAGGCGGCGGCTCAAGAGATCGGAGTGTGAGTTGGCCAG 1500
DB 1441 CGAGAGCGGAAAGGCTTAAGGCGGCGGCTCAAGAGATCGGAGTGTGAGTTGGCCAG 1500
QY 1501 GAGTGAAGTGGGCTGACCGGCGGCTGAGAGTGGCTGAGAGTGGCGGAGGAGAGCA 1560
DB 1501 GAGTGAAGTGGGCTGACCGGCGGCTGAGAGTGGCTGAGAGTGGCGGAGGAGAGCA 1560
QY 1561 CTTTGTCTTGTAGAGAGAGTACGAGGCTCCGAGGCTGAGAGAGAGAGAGAGTGC 1620
DB 1561 CTTTGTCTTGTAGAGAGAGTACGAGGCTCCGAGGCTGAGAGAGAGAGAGAGTGC 1620
QY 1621 CCGGCGAGAGTCTGAGCTGACGCGGCGGCTCGGCGGCTGAGAGAGAGAGAGTGC 1680
DB 1621 CCGGCGAGAGTCTGAGCTGACGCGGCGGCTCGGCGGCTGAGAGAGAGAGAGTGC 1680
QY 1681 GGGCGCGGAGCGGCTGAGAGTGGCTGAGAGTGGCTGAGAGTGGCGGAGAGAGCA 1740
DB 1681 GGGCGCGGAGCGGCTGAGAGTGGCTGAGAGTGGCTGAGAGTGGCGGAGAGAGCA 1740
QY 1741 CGTTCCGAGGCGTGTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
DB 1741 CGTTCCGAGGCGTGTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
QY 1801 GCCCAG 1806
DB 1801 GCCCAG 1806

Db 1801 GCCCAG 1806

RESULT 2
ABN95848
ID ABN95848 standard; DNA; 10884 BP.
XX
AC ABN95848;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #2346 used to diagnose liver cancer.
XX
KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
metastatic liver tumor; cytostatic; expression profile; disease state;
disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
PN WO200229103-A2.
XX
PD 11-APR-2002.
XX
PE 02-OCT-2001; 2001WO-US30589.
XX
PR 02-OCT-2000; 2000US-237054P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX
DR WPI; 2002-426119/45.
XX
PT Diagnosing and detecting the progression of liver cancer,
hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample -
XX
PS Claim 1; SEQ ID NO 2346; 298bp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumor in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 10884 BP; 2796 A; 2578 C; 2737 G; 2773 T; 0 other;

Query Match 99.4%; Score 1795; DB 24; Length 10884;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1806; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GGGGATTCGTTTGGTTCCTTCATTTTCCAAATGAAATCAGAGATCCTGTTGGGTCT 60
DB 101 GGGGATTCGTTTGGTTCCTTCATTTTCCAAATGAAATCAGAGATCCTGTTGGGTCT 160

QY 61 CAACCGAGTACTAGAGAGAGTGTATACAGAGAAAG- AAACGCAAGGAGCATATATG 119
DB 161 CAACGAGATTAAGAGAGAGTGTATACAGAGAAAG- AAACGCAAGGAGCATATATG 220

QY 120 CACGGTTTCCTTAACAAGGTGAGTGTAGCAGACGCTGAGCACTGTGGGAGAGAC 179
DB 221 CACGGTTTCCTTAACAAGGTGAGTGTAGCAGACGCTGAGCACTGTGGGAGAGAGC 280

QY 180 TCATAGAAAATAGCGGTGCTGGGCTTCCTCACCCCGGGCCCTCCATTTGTTCT 239
DB 281 TCATAGAAAATAGCGGTGCTGGGCTTCCTCACCCCGGGCCCTCCATTTGTTCT 340

QY 240 TTGGTCTCTTTTATTTTGTAGAGTCCAAATTTATTTATTTATTTAGTCAAGAGGAA 299
DB 341 TTGGTCTCTTTTATTTTGTAGAGTCCAAATTTATTTATTTATTTAGTCAAGAGGAA 400

QY 300 AATTGATCTTTCATCTTAAGAGAGTATATATGATATTAAGAGAGCTATATAT 359
DB 401 AATTGATCTTTCATCTTAAGAGAGTATATATGATATTAAGAGAGCTATATAT 460

QY 360 GGGGAGAGAGGTGAGACAGGGGGGAAAAGGGAGAGAGACAGAGAGAAAAGAGAGGAG 419
DB 461 GGGGAGAGAGGTGAGACAGGGGGGAAAAGGGAGAGAGACAGAGAGAAAAGAGAGGAG 520

QY 420 GGACAAAGAGAGACACTGGGCGAGAGATCGATTAGAGAGACAGAAATGATGATGAGA 479
DB 521 GGACAAAGAGAGACACTGGGCGAGAGATCGATTAGAGAGACAGAAATGATGATGAGA 580

QY 480 TTACCTTCACCCCAAGGCTTCGCTGGAGGGGAATGAGAGAGCTCTGATTTGCTATTA 539
DB 581 TTACCTTCACCCCAAGGCTTCGCTGGAGGGGAATGAGAGAGCTCTGATTTGCTATTA 640

QY 540 CTACTCCAAACTGCAAAAGGCTCTTCAAGTCACTATCCACCTCCTAAGGCAAGGCTCC 599
DB 641 CTACTCCAAACTGCAAAAGGCTCTTCAAGTCACTATCCACCTCCTAAGGCAAGGCTCC 700

QY 600 AATTTACAGCGCTGACGAAAGTCTCTCCGGGAGGTCTACCGCTTCCACTCCAC 659
DB 701 AATTTACAGCGCTGACGAAAGTCTCTCCGGGAGGTCTACCGCTTCCACTCCAC 760

QY 660 CCCCAAAACTCTTGGAAAAGTGCCTTGAAAATTTAATCCATTCATCCATCTGGAGC 719
DB 761 CCCCAAAACTCTTGGAAAAGTGCCTTGAAAATTTAATCCATTCATCTGGAGC 820

QY 720 ACCAGGCTCTCTGTGTGTCACCGAAGAGGGGGTGGCGAGCAAAACTGTAAGAACTCG 779
DB 821 ACCAGGCTCTCTGTGTGTCACCGAAGAGGGGGTGGCGAGCAAAACTGTAAGAACTCG 880

QY 780 AGTGCAGAGAGAGCCGACAGAGGTTACACCGACCTCAGGGCCATTTGGCCCCGAACT 839
DB 881 AGTGCAGAGAGAGCCGACAGAGGTTACACCGACCTCAGGGCCCAATTTGGCCCCGAACT 940

QY 840 TTACTGAAAAGTGTATAGATTGACAGATTAAGTATCCATCCAGCATGAGAAATACAG 899
DB 941 TTACTGAAAAGTGTATAGATTGACAGATTAAGTATCCATCCAGCATGAGAAATACAG 1000

QY 900 TAATACGAAGTGCCTTCAAAAATGACAATGAAAAATGCTATTAAAGACTATTTGGT 959
DB 1001 TAATACGAAGTGCCTTCAAAAATGACAATGAAAAATGCTATTAAAGACTATTTGGT 1060

QY 960 TAATACGTTTACAGAGTGCCTTCAGTTATTTGTTTATTTTGTGCTGGGTGTA 1019
DB 1061 TAATACGTTTACAGAGTGCCTTCAGTTATTTGTTTATTTTGTGCTGGGTGTA 1120

QY 1020 ACTCCATTTGAAAACATTAATCAGGGAGAAATCCCAAGACAGAAAGAGAGTGTCAATTA 1079
DB 1121 ACTCCATTTGAAAACATTAATCAGGGAGAAATCCCAAGACAGAAAGAGTGTCAATTA 1180

QY 1080 AATATATTTGAAAAGCCCTCCTTAAGAGACATTCGCTGCGGCTCCACTTTAATGGGG 1139
DB 1181 AATATATTTGAAAAGCCCTCCTTAAGAGACATTCGCTGCGGCTCCACTTTAATGGGG 1240

QY 1140 ACTTGGGTGTAGCAACACGTGAGATCTTTCTTGGCTTGAAGATTAAGCTGGAAGGCG 1199
DB 1241 ACTTGGGTGTAGCAACACGTGAGATCTTTCTTGGCTTGAAGATTAAGCTGGAAGGCG 1300

QY 1200 AAGGCCCCGGGCGATCTTGAATGAGTATTTGGGCGCCCTGGGGATTAACACAGCCGAG 1259
DB 1301 AAGGCCCCGGGCGATCTTGAATGAGTATTTGGGCGCCCTGGGGATTAACACAGCCGAG 1360

QY 1260 CGGGTGAATTAACCCGCAAGTGGCTGGCTCCAGACCCCAATGTAAGTCAGAAA 1319
 DB 1361 CGGGTGAATTAACCCGCAAGTGGCTGGCTCCAGACCCCAATGTAAGTCAGAAA 1420
 QY 1320 TGGCCCAAGCTTGGCTGGCAAGGAAATTAATTTAGAAAGTCCACGAGGTACAC 1379
 DB 1421 TGGCCCAAGCTTGGCTGGCAAGGAAATTAATTTAGAAAGTCCACGAGGTACAC 1480
 QY 1380 GAATGGGAGCGCTGATATGCAAGTTCACGACACCGGCTCGCCGCAAGGAGACCTCAC 1439
 DB 1481 GAATGGGAGCGCTGATATGCAAGTTCACGACACCGGCTCGCCGCAAGGAGACCTCAC 1540
 QY 1440 CCGAGAGCGGAAGGGTAAAGGGGGGGGTCAAGAGATCGGGGCTGCTAGTTGGCCA 1499
 DB 1541 CCGAGAGCGGAAGGGTAAAGGGGGGGGTCAAGAGATCGGGGCTGCTAGTTGGCCA 1600
 QY 1500 GGAGTACTGAGGAGTGAACGGGGGCTGAGAGTGGCTGAGAGTGGCGGGGTGGCGGAC 1559
 DB 1601 GGAGTACTGAGGAGTGAACGGGGGCTGAGAGTGGCTGAGAGTGGCGGGGTGGCGGAC 1660
 QY 1560 ACCTTGGTCTTGTAGACGAACAAGTGAACGGGCTCGGGGCTGAGCAAGAGACAGGTG 1619
 DB 1661 ACCTTGGTCTTGTAGACGAACAAGTGAACGGGCTCGGGGCTGAGCAAGAGACAGGTG 1720
 QY 1620 CCGGGGCGAGTCTGAGACTGACAGCCCGAGAGTCCCGGCTGCTCAGGGCGAAACA 1679
 DB 1721 CCGGGGCGAGTCTGAGACTGACAGCCCGAGAGTCCCGGCTGCTCAGGGCGAAACA 1780
 QY 1680 CCGGGGCGAGTCTGAGACTGACAGCCCGAGAGTCCCGGCTGCTCAGGGCGAAACA 1739
 DB 1781 CCGGGGCGAGTCTGAGACTGACAGCCCGAGAGTCCCGGCTGCTCAGGGCGAAACA 1840
 QY 1740 GCGTTCCGGGGCGTGTCTGAGGCAACCCCGCTTATATATACGGGCGCGCGGCG 1799
 DB 1841 GCGTTCCGGGGCGTGTCTGAGGCAACCCCGCTTATATATACGGGCGCGCGGCG 1900
 QY 1800 CGCCGAG 1806
 DB 1901 CGCCGAG 1907
 RESULT 3
 ABL33002
 ID ABL33002 standard; DNA: 8700 BP.
 AC ABL33002;
 DT 26-MAR-2002 (first entry)
 XX
 DE Human immune system associated gene SEQ ID NO: 975.
 XX
 KW Human; immune system disease; cytosine methylation; antiautismic;
 KW antileukemic; antianemic; cytosine; cytosine; cytosine; cytosine;
 KW antileukemic; antianemic; cytosine; cytosine; cytosine; cytosine;
 KW antileukemic; antianemic; cytosine; cytosine; cytosine; cytosine;
 KW antileukemic; antianemic; cytosine; cytosine; cytosine; cytosine;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200200928-A2.
 PD 03-JAN-2002.
 XX
 PF 02-JUL-2001; 2001WO-EP07537.
 XX
 PR 30-JUN-2000; 2000DE-103529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPIC-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K;
 XX WPI: 2002-130909/17.
 DR
 XX
 PT Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation
 PT
 XX
 PS Claim 1; SEQ ID NO 975; 32bp + Sequence Listing; German.
 CC
 CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 CC
 XX
 SQ Sequence 8700 BP; 2247 A; 281 C; 223 G; 3939 T; 0 other;

Query Match 63.2%; Score 1141.4; DB 24; Length 8700;
 Best Local Similarity 80.0%; Pred. No. 1.2e-294;
 Matches 1454; Conservative 0; Mismatches 351; Indels 12; Gaps 9;

QY 1 GGGGATTCGTTTGTCTTCAATTTCCAAATGAAATCAGATCTGTTTGGGTGT 60
 DB 3055 GGGGATTCGTTTGTCTTCAATTTCCAAATGAAATCAGATCTGTTTGGGTGT 3114
 QY 61 CAACGCAATCTGAAGAGGATGATCAAGAGAAAGG-AACGCAAGCAGATTAATG 119
 DB 3115 TAAAGTATATTAAGAGAGGATGATCAAGAGAAAGG-AACGCAAGCAGATTAATG 3174
 QY 120 CACGTTTCTGTAACAAAGTGAAGTGAAGCAGACCTGAGCAGTGGGAGAGAGC 179
 DB 3175 TACGTTTCTGTAACAAAGTGAAGTGAAGCAGACCTGAGCAGTGGGAGAGAGC 3234
 QY 180 TCATAGAAATGAGGAGTGTGCTGCTCAACCCCGGGGCTTCATTTGTTGCT 239
 DB 3235 TTATAGAAATGAGGAGTGTGCTGCTCAACCCCGGGGCTTCATTTGTTGCT 3294
 QY 240 TTGGTCTCTTTATTTGATGAGTCCATTTATTTATTTAGTACAAAGAGAGACA 299
 DB 3295 TTGGTCTCTTTATTTGATGAGTCCATTTATTTATTTAGTACAAAGAGAGACA 3354
 QY 300 AATTGATCTTTCATTTCAATTAAGAGAGATATATATTAAGAGAGAGAGAGAG 359
 DB 3355 AATTGATCTTTCATTTCAATTAAGAGAGATATATATTAAGAGAGAGAGAGAG 3414
 QY 360 GGGGGAAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 419
 DB 3415 GGGGGAAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3474
 QY 420 GGACAG 479
 DB 3475 GGATTAAG 3534
 QY 480 TTAATCTCACCAAGGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 539
 DB 3535 TTAATCTCACCAAGGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3593
 QY 540 CTACTCCAAATGCAAGAGGCTCTCAAGTCACTATCCATCCCTAAGGAGAGGTC 599
 DB 3594 TTAATCTCACCAAGGCTCTCAAGTCACTATCCATCCCTAAGGAGAGGTC 3653
 QY 600 AATTCAACAGCGTTCAGAGAGAGTCTCCGAGAGAGTTCACCGCTCCACTCAC 659
 DB 3654 AATTCAACAGCGTTCAGAGAGAGTCTCCGAGAGAGTTCACCGCTCCACTCAC 3713
 QY 660 CCCACAACTCTTGAAGAGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 719
 DB 3714 TTTTATTAATTTTGGAGAGAGGTTTGAGAGAGAGAGAGAGAGAGAGAGAGAG 3773

QY 720 ACCAGGCTCTCTGTGTCACCGAA -GGAGGGGTCGCCAGACAAACTGAAAGAACTC 778
 Db 3774 ATTAGGCTTTTGTGTTATGATGAGGAGGGGGTCCGATGATGAAATGAAAGAAATTC 3833
 QY 779 GAGTGCAGAGAAAGCCGACAGAGATTACAGCACTCAGCCGCAATTCG -CCCCGAA 837
 Db 3834 GAGTGTAGAGAAAGGTCGATAGAGATTATGCAATTTTACGGCTAATTCGTTTCGAA 3893
 QY 838 CTTTACTGAAAGTGTGTAGATTGACAGATTAAGTAAATCCCAACGATCGAGAAATAC 897
 Db 3894 TTTTATGAAAGTGTGTAGATTGATGATGATGATGATGATGATGATGATGATGATGAT 3953
 QY 898 AGTAATACGAAAGTCCCTCAAAATGACATGAAATTCCTATTAAAGCATATTG 957
 Db 3954 AGTAATACGAAAGTCCCTCAAAATGACATGAAATTCCTATTAAAGCATATTG 4013
 QY 958 GTTAAATACGTTTACAGAGCCAGTTTATGCTTATTTTATTTTTCGAGGCTG 1017
 Db 4014 GTTAAATACGTTTATGAGTGTGTTATGTTTATTTTATTTTATTTTTCGAGGCTG 4073
 QY 1018 AAATCCATTTGAAACATTAATCAGGAGATTAACCAAGACAAAGAAAGATGTCAT 1077
 Db 4074 AAATTTATTTGAAATTAATTAAGAGAAATTAATTAAGATTAAGAAATTAATTTAT 4133
 QY 1078 TAAATATTTGAAAGCCCTGCTTAAGGA -GATTCGCTGCGCGCTCCACTTATTTG 1136
 Db 4134 TAAATATTTGAAAGCCCTGCTTAAGGA -GATTCGCTGCGCGCTCCACTTATTTG 4193
 QY 1137 GGGACTTGGGCTGACACACGTCAGAGAGCTCTTGGCTGAGAAAGTAAGTAAGTAAG 1196
 Db 4194 GAGATTTGGCGTGTAGTAATACGTGAGAGTTTTCGCTGAGAAAGTAAGTTGGAAG 4253
 QY 1197 GCG - -AAGGCCCCGGGATCTTCAAGATGCTATTTGAGGAGCCCTGGGATTAACA 1253
 Db 4254 GGGCAAGGTTTGGGCGATTTTATTTAGATGCTATTTGAGGAGTTTGGGATTAACA 4313
 QY 1254 GCCCAGCGGTGTAAATTAACCCGCACTGCTGCTCCGAGAACCAATTTAGT 1313
 Db 4314 GTTACCGGGTGTAAATTTAAATTTCTAGGTTTGGTTTAAATTTAAATTTAAATTT 4373
 QY 1314 CAGAAATGTCACAGATTCGCTGCAACGGAATTAATTTAGAAAGTCCACGAGT 1373
 Db 4374 TAGAAATGTTTAAATTTCTGTTGTTTACGAAATTAATTTAGAAAGTTTACGAGG 4433
 QY 1374 ACACAGAAATGCGAGCGCTGTATGCAATTTCCCGACACGCGCTGCGCGAGAGAC 1433
 Db 4434 ATATAGAAATGCGAGCGCTGTATGCAATTTCCCGATATGCGTTCTGTAAGGAGAT 4493
 QY 1434 CTCACCCCGAGAGCGGAAGGGGTAAAGGCGCGGGTCAAGAGATCGGGGCTGAGT 1493
 Db 4494 TTTATTTGAGAGCGGGAAGGGGTAAAGGCGCGGGGTAAAGAGATCGGGGCTGAGT 4553
 QY 1494 TGGCCAGAGATGATGAGGGGATCCGGGGGTGCTGAGGTGCGCGGGGGTGGCC 1553
 Db 4554 TGGTTAGAGATGATGAGGGGATCCGGGGGTGCTGAGGTGCGCGGGGGTGGTC 4613
 QY 1554 GGGCACACCTTGGTCTGTAGACGACAAAGTGA -CGGCTTCGGGCGGTGAGACAGAG 1611
 Db 4614 GGGTATATTTTGGTTTGTAGACGATTAAGGTGATTCGGGTTTCGGGCGGTGAGAGG 4673
 QY 1612 AGCAGTGGCCCGGAGAGTCTGAGCTGACGCCGCCGAGCTGGCCCGGCTGCTAG 1671
 Db 4674 AGTAGTGTGTCGGGAGGTTGAGTGTGATGCTTTCGAGTTGCTGTTTCTGTTTAA 4733
 QY 1672 GCGAAGCAGGAG -CCCCGAGCGCGCTGCGCGGAGCCGCCGCCGCCCAACCCCACT 1730
 Db 4734 GCGAATACGAGGTTTTCGATGCTGTTTTCGCGGATGCTTTTATTTTATTTTATTT 4793
 QY 1731 CCGTGGCGCG -CGTTCGGGCGGTCTCTGCGGCGCACCCGCGCTTATATAGCGCGCG 1789
 Db 4794 TTTGGCGCGGTGCTTTCGGGCGGTGTTTGGGTTTATTTGCGTTTATATAGCGCGCG 4853
 QY 1790 GCGCGCGCGCGCGCGCG 1806

Db 4854 GCGTTCGGGCTGTTAG 4870
 RESULT 4
 ABL33003/C
 ID ABL33003 standard; DNA: 8700 BP.
 ABL33003:
 26-MAR-2002 (first entry)
 Human immune system associated gene SEQ ID NO: 976.
 DE Human immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; antianaemic; cytosolic; neotrophic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antineumatic; antiarthritic; antidiabetic; antipsoriatic;
 KW antinflamatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.
 XX Homo sapiens.
 OS MO200200928-A2.
 PN 03-JAN-2002.
 PD 02-JUL-2001: 2001MO-EP07537.
 PF 30-JUN-2000: 2000DE-1032529.
 PR 01-SEP-2000: 2000DE-1043826.
 PA (EPIC-) EPIGENOMICS AG.
 XX Olek A, Piepenbrock C, Berlin K;
 PI WPI; 2002-130909/17.
 DR Nucleic acid compitling fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation -
 XX
 PS Claim 1; SEQ ID NO 976; 32pp + sequence listing; German.
 CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 XX
 SQ Sequence 8700 BP; 2133 A; 281 C; 2087 G; 4199 T; 0 other;
 Query Match 56.5%; Score 1021.2; DB 24; Length 8700;
 Best Local Similarity 76.0%; Pred. No. 1.8e-262;
 Matches 1377; Conservative 0; Mismatches 423; Indels 12; Gaps 9;
 QY 5 ATTCGTTTGTGTTTCTTCAATTTTCAATGAAATCAGAGATCTGTTGCGTCAAC 64
 Db 5642 ATTCGTTTATTTCTTCAATTTTCAATGAAATCAGAGATCTGTTGCGTCAAC 5583
 QY 65 GCAGATACAGAGAGGATACAGAGAGAAAGAA -CAGCAAGCAGATTTGCGAGC 123
 Db 5582 GCAGAAATACAGAGAGGATACAGAGAGAAAGAA -CAGCAAGCAGATTTGCGAGC 5523
 QY 124 GTTTCGTTAAAGAGTGTAGTGCACACCGTACAGACGTTGGGAGAAAGCGTCAT 183
 Db 5522 ATTTCTATTAACAAATTAATTAATTAACCAACCTTAACACGTTATTAATAAATTAAT 5463

OY	184	AGAAATATACGGTGTGGGCTTCGTCACCCGGGGCCCTCCATTTGCTCTTG	243
Db	5462	AAAAAATATACGATACTAAACCTTCGTCACCCGAAACACTCCATTTCTATCTTAA	5403
OY	244	TCCTCTTTTATTTGTAGAGGTCCAATTATTTATTTATGTACAGAGGACGAATTT	303
Db	5402	TCCTCTTTTATTTATTAATAATCCAAATTTATTTATTTATTAATCAAAAAAGAAATTT	5434
OY	304	GATCTTTCAATCTATAAGAGGTATATTTGTATTAAGGAACCTATATGATATGGG	363
Db	5342	AATCTTTCAATCTATAAAAAAATATATATATTAATAAAAAAACTATATTAATATATAA	5283
OY	364	GAAGAGGTGAGCGGGGAAAGAGGGGAGAGAGAGAGAAAGAGAGAGAGAGAC	423
Db	5282	AAAAAATATTAACAAAAAATTAATAATAATAATAATAATAATAATAATAATAATAA	5223
OY	424	AAGAGAGACACTGGCGAGAGATCGATTTAGAGAGACAGAAATGATGAATTA	483
Db	5222	AAAAAAAAAACTTAACGAAAAATTCGATTAATAAAAAACAAAAAATTAATAATAATAA	5163
OY	484	CTTCAACCCCAAGCTTGCTGCTTGAGAGGGAATGAGAGACTCCGATTTGCTATTACTAC	543
Db	5162	CTTCAACCCCAAACTTTGCT-ACATAAAAAAATATAAAAAATCTCAATTTCTATTACTAC	5104
OY	544	TCCAAACTGCAAAAGGGCTCTTCAAGTCACTTTCCACTCTCTTAAGGCAAGCGTCATTT	603
Db	5103	TCCAAACTGCAAAAAAATCTCTTAAATCACTTTCCACTCTCTTAAGCAAAACGTCATTT	5044
OY	604	TCACAGCGCTTGAGAAAGTCTCTCCCGGAGGTCCTACCGCTTCCCATCTCACCCCC	663
Db	5043	TCACAGCGCTTCAAAAAATCTCTCCCGGAGAAATCTCACCGGTTCCCATCTCACCCCC	4984
OY	664	ACAAACTCTTTGAAAAAGTGCCTTGAATAATTTATTCCTCAATCCAAATCCTGGACACCA	723
Db	4983	ACAAACTCTTTAAAAAATACCTTAAAAAATTTATTCCTCAATCCAAATCCTAAACACCA	4924
OY	724	GGCTCTCTGTGGTGCACGG-AAAGGAGGGGTGGCGAGCAAAAACTGAAGAACTCGAGT	782
Db	4923	ACGCTCTCTATTATATACCGGAAAAAATAATACGCAACAAATCTAAAAAACTCGAAT	4864
OY	783	GGCAGAGAGGCGGACAGAGATTACAGCGACTCAGCGGCGCAATTGGC-CCCGCACTTT	841
Db	4863	ACCAAAAAAAGCGACAAAAATTTACACGACTCAACGGCGCAATTACGCCCCGCACTTT	4804
OY	842	ACTGAAAGACTGTTTAGATTGCGACAGATTAAGCTTGAATCCCAAGCATCGAGATTCAGTA	901
Db	4803	ACTAAAAAATATTTAAATTACAAAAATTAACATAAATCCCAAGCATCGAAAAATACATA	4744
OY	902	ATACGAGTGGCTTCAAAAAATGACAAATGAATAATGGCTATTTAAAGCACTATTGGTTA	961
Db	4743	ATACGAAATTCGCTTCAAAAAAATTAACAAATTAATAATTAATAAAACCTATTTAATTA	4664
OY	962	ATTAGGTTTCAGCAGTGGCCAGTTTATGTCTTATTTATCTTTTGCTGGGTGTAAAC	1021
Db	4663	ATTAGGTTTCAGCAATTAACCAATTTATTTATTTATTTATTTATTTATTCGTAATATTAAC	4624
OY	1022	TCCATTGAAAAATATATCAGGAGATTAACCAAGACAGAGACAGTGTCTATTTAA	1081
Db	4623	TCCATTGAAAAATATATCAAAAAAATATACCCCAAAAAACAAATATATCTTTAA	4564
OY	1082	ATATTGAAAAAGCCCGCTTTAAGA-GCATTTGGCTTGCCGTCACACTCTTAATTTGGGGA	1140
Db	4563	ATATTGAAAAAGCCCTTATTAAGCAATTTGCTTTTACCTCGATCTCTTAATTTAAAA	4504
OY	1141	CTTGGCGTGTAGCAACACGAGTAGAGTCTTCTGCTTGAAGATGAACCT--GGAAGG	1197
Db	4503	CTTAGCATTTAACACACGCTAAAAATCTTTTACGTTTAAAAAATTAACCTTAATAAAC	4444
OY	1198	CGAAGCCCGGGGCATCTTTCAGATCGATTTTGTGGGCCCTGGGGATTTAAACAGCC	1257
Db	4443	GAAACCCCGAAGCATCTTCAAAATACGATTTATTAACCCCTAAAAAATTAACAAACCC	4384
OY	1258	AGCGGCTGTAATTAACCCCGCATGCTTGGCTCCCTGAGACCCAAATGTATGTCGA	1317

Db	4383	AACGAATATAATTAAACCCCGCAATACCTTAATCTCCATAAAACCCAAATATAATTAATCAA	4328
Qy	1318	AATGTCCCAAGACTTCCGCTGCCCACAGGAATTAAATTTTGAAGACTCCACAGAGTACAC	1377
Db	4333	AATATCCCAAAACTTGCCTCAACCAACAAATTAATTTTAAAAAACTCCACAAAATATAC	4264
Qy	1378	ACGAATGGGAGCGCTGTATGCGCAGTTTCCCGACACCGGCTGGCGCGAGGAGACTCA	1437
Db	4253	ACGAATAGAAACCTTATATACCAATTTTCCCGACACCGACTGGCGCAAAAAACCTCA	4204
Qy	1438	CCCCGAGACGGGAAGGGGTAAAGGGCGGGGGTCAAGAGATCGGGGGTGTGACTGTGC	1497
Db	4203	CCCCGAAAACGAAAAAATATAAAAACGACAAATCAAAAAATCGAAAAATATTAATTAAC	4144
Qy	1498	CAGAGTAGCTGGGGGTACCGGGGGTGTCTGAGTGCGCTGAGTGCCTGGGGTGGCGGGG	1557
Db	4143	CAAAATATACATAAAATAAACGAAAATATCTAAATTAATACCTAAATATCCGAAATATACGAAAC	4084
Qy	1558	ACACCTTGTTCTTGTAGAGACACAAAGGTGA--CGGGCTCCGGGGCTGAGACACGAGAGACA	1615
Db	4083	ACACCTTAATTTCTTATTAAGACAAAATTAACCGGAATCGGAACCTCGAAGCTAGCCACAGAAAACA	4024
Qy	1616	GGTGCCCCGGGGGAGTCTTCGAGCTGACAGCCGCCCGGACTGGGCCCGGGCTGCTCAGAGCGGA	1675
Db	4023	AATACCCGGAAGATCTTCGAACTACACAGCCCCCGGAATCGAACCCGCACTACTCAAAACGA	3964
Qy	1676	AGCAGC--GGCCCCGACCCGCTGCGCTGGGGCGAGCCGCCGCCCTGCCAACCCCTCACTCCG	1734
Db	3963	AACACGAACCCCGGCACACTACTACTAGCCGACACCGGCCCTCCCTCCACCCCCCACTCCTA	3904
Qy	1735	GGGCGC--CGTTCCGGGGCGTGTCTCTGGGCCACCCCGGCTTCTATATACGGGCGGGCGGC	1793
Db	3903	AACCGCGGTTCCGGAACAGATATCTAAACACACCCGCACTTATATTAACGACCGGACGCGC	3844
Qy	1794	CCGGGCGGCCCA 1805	
Db	3843	CCGAACGCGCCA 3832	

XX	RESULT 5
ABQ45156/C	
ID	ABQ45156 standard; DNA; 1007 BP.
XX	
AC	ABQ45156;
XX	
DT	12-JUL-2002 (first entry)
DE	
XX	Oligonucleotide for detecting cytosine methylation SEQ ID NO 31747.
XX	
XX	Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW	drug; side effect; cancer; central nervous system; cardiovascular;
KW	gastrointestinal; respiratory system; single nucleotide polymorphism;
KW	SNP; cell differentiation; ds.
XX	
OS	Homo sapiens.
XX	
PN	W0200218632-A2.
XX	
PD	07-MAR-2002.
XX	
XX	
PF	01-SEP-2001; 2001WO-EP10074.
XX	
PR	01-SEP-2000; 2000DE-1043826.
PR	05-SEP-2000; 2000DE-1044543.
XX	
PA	(EPig-) EPIGENOMICS AG.
XX	
PI	Olek A, Piepenbrock C, Berlin K, Guetig D;
XX	
DR	WPI; 2002-371829/40.
XX	
XX	Determining the degree of cytosine methylation in genomic DNA, useful

Db 889 GAAAGAA 883

RESULT 12
ABN81171

ID ABN81171 standard; DNA; 170 BP.

XX ABN81171;

DE 16-JUL-2002 (first entry)

XX Shrimp polynucleotide TUDGLV3-5.289F.

XX Giant black tiger prawn; Penaeus monodon; pacific white shrimp;

XX Litopenaeus vannamei; shrimp; microsatellite sequence; genome mapping;

XX Taura Syndrome Virus; TSV; infection; ds.

OS Litopenaeus vannamei.

PN WO200034476-A2.

PD 15-JUN-2000.

XX 10-DEC-1999; 99WO-US29571.

XX 10-DEC-1998; 98US-0111670.

XX (TUFTS) TUFTS COLLEGE.

PI Alciwar-Warren A, Xu Z, Dhar AK, Fan Y, Meehan D, Garcia DK;

DR WPI; 2000-423422/36.

XX Polynucleotides of shrimp are useful for identifying, mapping and

PS Claim 1, Fig 1; 120pp; English.

XX The invention relates to an isolated polynucleotide (1) of the giant

CC black tiger prawn, Penaeus monodon or expressed sequence tags of the

CC pacific white shrimp, Litopenaeus vannamei (ABN80997-ABN81172), both

CC containing microsatellites sequences including those p. monodon

CC complementary sequences given in Genbank AF077550-AF077598. (1), the

CC for mapping of the genome of various species of shrimp. Mapping the

CC genome of Penaeus is useful for determining whether a test shrimp,

CC preferably Litopenaeus vannamei, has a genotype associated with a

CC phenotypic trait such as resistance to Taura Syndrome Virus (TSV)

CC infection.

XX Sequence 170 BP; 78 A; 11 C; 73 G; 8 T; 0 other;

XX

XX

XX

XX

XX

XX

XX

XX

XX

AC ABN81079;

XX 16-JUL-2002 (first entry)

DE Shrimp polynucleotide SEQ ID NO 83.

XX Giant black tiger prawn; Penaeus monodon; pacific white shrimp;

XX Litopenaeus vannamei; shrimp; microsatellite sequence; genome mapping;

XX Taura Syndrome Virus; TSV; infection; ds.

OS Penaeus monodon.

PN WO200034476-A2.

PD 15-JUN-2000.

XX 10-DEC-1999; 99WO-US29571.

XX 10-DEC-1998; 98US-0111670.

XX (TUFTS) TUFTS COLLEGE.

PI Alciwar-Warren A, Xu Z, Dhar AK, Fan Y, Meehan D, Garcia DK;

DR WPI; 2000-423422/36.

XX Polynucleotides of shrimp are useful for identifying, mapping and

PS Claim 1, Page 97; 120pp; English.

XX The invention relates to an isolated polynucleotide (1) of the giant

CC black tiger prawn, Penaeus monodon or expressed sequence tags of the

CC pacific white shrimp, Litopenaeus vannamei (ABN80997-ABN81172), both

CC containing microsatellites sequences including those p. monodon

CC complementary sequences given in Genbank AF077550-AF077598. (1), the

CC for mapping of the genome of various species of shrimp. Mapping the

CC genome of Penaeus is useful for determining whether a test shrimp,

CC preferably Litopenaeus vannamei, has a genotype associated with a

CC phenotypic trait such as resistance to Taura Syndrome Virus (TSV)

CC infection.

XX Sequence 502 BP; 81 A; 162 C; 41 G; 218 T; 0 other;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

AC ABN81079;

XX 16-JUL-2002 (first entry)

DE Shrimp polynucleotide SEQ ID NO 83.

XX Giant black tiger prawn; Penaeus monodon; pacific white shrimp;

XX Litopenaeus vannamei; shrimp; microsatellite sequence; genome mapping;

XX Taura Syndrome Virus; TSV; infection; ds.

OS Penaeus monodon.

PN WO200034476-A2.

PD 15-JUN-2000.

XX 10-DEC-1999; 99WO-US29571.

XX 10-DEC-1998; 98US-0111670.

XX (TUFTS) TUFTS COLLEGE.

PI Alciwar-Warren A, Xu Z, Dhar AK, Fan Y, Meehan D, Garcia DK;

DR WPI; 2000-423422/36.

XX Polynucleotides of shrimp are useful for identifying, mapping and

PS Claim 1, Page 97; 120pp; English.

XX The invention relates to an isolated polynucleotide (1) of the giant

CC black tiger prawn, Penaeus monodon or expressed sequence tags of the

CC pacific white shrimp, Litopenaeus vannamei (ABN80997-ABN81172), both

CC containing microsatellites sequences including those p. monodon

CC complementary sequences given in Genbank AF077550-AF077598. (1), the

CC for mapping of the genome of various species of shrimp. Mapping the

CC genome of Penaeus is useful for determining whether a test shrimp,

CC preferably Litopenaeus vannamei, has a genotype associated with a

CC phenotypic trait such as resistance to Taura Syndrome Virus (TSV)

CC infection.

XX Sequence 502 BP; 81 A; 162 C; 41 G; 218 T; 0 other;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 23, 2003, 08:21:38 ; Search time 2375 Seconds
(without alignments)
12315.389 Million cell updates/sec

Title: US-09-865-879-1

Perfect score: 1806
Sequence: 1 gggagatcgtttgttccct.....ggcgccgcggcgccgccag 1806

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estba:*
2: em_estchum:*
3: em_estlin:*
4: em_estlinu:*
5: em_estlov:*
6: em_estlovl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl1:*
10: gb_estl2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81.6	4.5	485	17	BH302763
2	74.4	4.1	1203	17	BH302763 CH230-970
3	71.6	4.0	932	17	CNS0015Y4
4	70.8	3.9	1364	14	AL066742 Drosophila
5	70.6	3.9	1171	14	BM810045 AGENCOURT
6	70.2	3.9	1160	14	BM914930 AGENCOURT
					BM926731 AGENCOURT

C	7	69.8	3.9	1021	14	BM434910	BM434910 AGENCOURT
C	8	69.2	3.8	1273	13	BM562099	BM562099 AGENCOURT
C	9	68.8	3.8	935	17	CNS006XK	AL066051 Drosophila
C	10	67.6	3.7	1190	14	BM673485	BM673485 AGENCOURT
C	11	67.2	3.7	839	17	CNS004NB	AL054280 Drosophila
C	12	67.2	3.7	925	17	CNS0091P	AL053013 Drosophila
C	13	67	3.7	1016	14	BM18616	BM18616 AGENCOURT
C	14	66.8	3.7	1162	14	BM051092	BM051092 AGENCOURT
C	15	66.6	3.7	1136	14	BM0943816	BM0943816 AGENCOURT
C	16	66	3.7	972	13	BM561468	BM561468 AGENCOURT
C	17	66	3.7	995	14	BM039901	BM039901 AGENCOURT
C	18	66	3.7	1169	13	BM549879	BM549879 AGENCOURT
C	19	65.6	3.6	233	14	BM0556810	BM0556810 H4042C11-
C	20	65.6	3.6	1315	13	BM543062	BM543062 AGENCOURT
C	21	65.4	3.6	927	14	BM680645	BM680645 AGENCOURT
C	22	65.2	3.6	452	17	BM740100	BM740100 gtf2a03.b
C	23	65.2	3.6	845	17	BM032983	BM032983 Pan tlog1
C	24	65	3.6	987	14	BM055647	BM055647 AGENCOURT
C	25	65	3.6	997	14	BM053024	BM053024 AGENCOURT
C	26	65	3.6	1138	14	BM070284	BM070284 AGENCOURT
C	27	64.4	3.6	1417	14	BM810236	BM810236 AGENCOURT
C	28	64.4	3.6	467	17	AL235878	AL235878 Tetradon
C	29	64.4	3.6	471	17	AZ066055	AZ066055 RPT-23-4
C	30	64.4	3.6	939	17	AG043613	AG043613 Pan tlog1
C	31	64.4	3.6	1302	14	BM061839	BM061839 AGENCOURT
C	32	64	3.5	1176	13	BM544862	BM544862 AGENCOURT
C	33	63.8	3.5	840	17	AG043467	AG043467 Pan tlog1
C	34	63.8	3.5	1101	17	AL108773	AL108773 Drosophila
C	35	63.6	3.5	888	12	BM809572	BM809572 mgcl001a
C	36	63.4	3.5	573	13	BM194158	BM194158 sc83a03.y
C	37	63.4	3.5	925	17	CNS0091P	AL053013 Drosophila
C	38	63.2	3.5	788	17	AG032933	AG032933 Pan tlog1
C	39	63.2	3.5	1269	14	BM807701	BM807701 AGENCOURT
C	40	63.2	3.5	1315	13	BM543062	BM543062 AGENCOURT
C	41	62.6	3.5	531	14	BM051546	BM051546 EST622861
C	42	62.6	3.5	967	14	BM078596	BM078596 AGENCOURT
C	43	62.6	3.5	1101	17	CNS00397	AL063912 Drosophila
C	44	62.4	3.5	374	12	BM087347	BM087347 H3138E08-
C	45	62.4	3.5	466	17	AZ040655	AZ040655 RPT-23-3

ALIGNMENTS

RESULT 1	BH302763	485 bp	DNA	linear	GSS 30-NOV-2001
LOCUS	CH230-97018	TV	CHORI-230	Segment 1	Rattus norvegicus genomic clone
DEFINITION	CH230-97018	DNA sequence.			
ACCESSION	BH302763				
VERSION	BH302763.1	GI:17215171			
KEYWORDS	GSS...				
SOURCE	Norway rat.				
ORGANISM	Rattus norvegicus				
REFERENCE	1 (bases 1 to 485)				
AUTHORS	Zhao,S., Shetty,T., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de Jong,P. and Fraser,C.M.				
TITLE	Rat BAC End Sequences from Library CHORI-230 EcoRI segment				
JOURNAL	Unpublished (1999)				
COMMENT	Other GSSs: CH230-97018.TJ				
	Contact: Shanying Zhao				
	Department of Eukaryotic Genomics				
	The Institute for Genomic Research				
	9712 Medical Center Dr., Rockville, MD 20850, USA				
	Tel: 301 838 0200				
	Fax: 301 838 0208				
	Email: szhaoc@igf.org				
	Clones are derived from the rat BAC library CHORI-230				
	(http://www.chori.org/bacpac/rat230.htm). For BAC library				

[illegible]

RESULT 9	CNS006XK	935 bp	DNA	linear	GSS 03-JUN-1999
CNS006XK					
LOCUS					
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC #				
	BACR14N09 of RP11-98 library from Drosophila melanogaster (fruit				
	fly), genomic survey sequence.				
ACCESSION	AL066051				
VERSION	AL066051.1				
KEYWORDS	GI:4945019				
SOURCE	GSS.				
ORGANISM	Drosophila melanogaster.				
	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peleiygota;				
	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
	Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 935)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-JUN-1999)				
	Genoscope - Centre National de Sequencage :				

COMMENT

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Oseegawa and Aaron Mammasser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCL-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES	Location/qualifiers
SOURCE	1. .935

	BASE COUNT	BRIGIN
/organism="Drosophila melanogaster"		
/db_xref="taxon:7227"		
/clone="BACRI4N09"		
/clone_1b="RRCI-98"		
/note="end : T7"		
257 a	170 c	162 g 96 t 250 others

[illegible]

QY	1501	GAGTGCATGGGAGTGACCGGAGGTGGCTGAGTGGTCCCTGGAGTGGAGCGGGGTGGCCGAGACA	1560
Db	623	SGCCGCGCGSGGCGCGGCGSSGSGSSGSSGSSGSSGSSGSSGSSGSGGCGGCGGCGCGC	682
QY	1561	CCATTGATTCTTGTAGACGACCAAGGTGACG - GGCTCCGCGCGCTGAGCAGCAGAGACAGT	1618
Db	683	GCSCGSCGSCGCGSSGCGSGCGSGCGSGGAGCGGSCGCSGCGSCGSGSGSGSGSCGSC	742
QY	1619	GCCCGGCGGAGTCTCGAGCTGCACGCGCCCGGAGCTCGGCGCCCGCGCTGCTCAGGGCGCAAGC	1678
Db	743	GSCTSSSSSSSCSSCGCGCGCMSCSGCGGSSCGCGCCGCCGCGCGCGGCGCGCGC	802
QY	1679	ACGGGGCCCCGACCGCGTGTGCGCGGACCGCGCCCGCCCTCCCAACCCCGCACTCTGGGGG	1738
Db	803	GCGGSGCGGCGCGCGSSGGGSGCGCGCGCGCGCCSSCGAGCGGSGGSSGCG	862
QY	1739	CGCGTTCGCGGAGGCGTGACCTGAGGCGCACCCCGGCTTCTATTATACGGGCGGCGCGCCGGG	1798
Db	863	SGGCGSGGSGGSSGSGSSGSGSSGSGGCGSGGAGGSGGAGGAGCCGCGCGCGCGSSSS	922
QY	1799	CGCGCC 1804	
Db	923	SCGSCC 928	

RESULT 10	
B0673485	
LOCUS	B0673485
DEFINITION	AG0673485 1190 bp mRNA linear EST 15-JUL-2002
ACCESSION	AB0673485.1 GI:21784319 Homo sapiens cDNA IMAGE:6254865
VERSION	B0673485.1
KEYWORDS	5' UTR sequence.
SOURCE	human.

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1190)				
NIH-MGC http://mgc.nci.nih.gov/ .				
National Institutes of Health, Mammalian Gene Collection (MGC) unpublished (1999)				
Contact: Robert Strausberg, Ph.D.				

Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINT)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MSC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINT at:
<http://image.llnl.gov>
 Plate: L10CM2405 row: a column: 10
 High quality sequence stop: 226.
 Location/Qualifiers

```

source
1. 1190
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1 IMAGE:6254865"
/clone_11b="NIH_MGC_102"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: salivary gland; Vector: pOT57; Site_1: XhoI; Site_2: EcoRI; cDNA made by 0190-ct priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCCACGAC(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)." Note: this is a NIH_MGC Library."
30 a 471 c 581 g 69 t 39 others

```

Query Match	3.7%;	Score 67.6;	DB 14;	Length 1190;
Best Local Similarity	50.6%;	Pred. No. 5.8e-05;		

RESULT 11	CNS004NB	LOCUS	DEFINITION
	CNS004NB	839 bp	DNA
			linear GSS 03-JUN-1995
			Drosophila melanogaster genome survey sequence TR13 end of BAC #
			BA010161 of R101-98 library from Drosophila melanogaster (fruit
			fly), genomic survey sequence.

ACCESSION	AL054280
VERSION	AL054280.1
KEYWORDS	GI:4931788
SOURCE	GSS.
ORGANISM	Drosophila melanogaster.
	Drosophila melanogaster.
	Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
	Ephydroidea; Drosophilidae; Drosophila.
REFERENCE	1 (bases 1 to 839)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage

COMMENT

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genome.cns.fr)
Web : www.genome.cns.fr

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see [http://www.fruitfly.org/The BDGP Drosophila melanogaster BAC library](http://www.fruitfly.org/The%20BDGP%20Drosophila%20melanogaster%20BAC%20library) was prepared by Kazuhiro Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPci-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2: cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

1.839

```

/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_1b="BACR10E16"
/clone_lib="RPCT-98"
/note="end : TET3"

```

BASE C
ORIGIN

Query Match

3.78; Score 67.2; DB 17; Length 839,

Best Local Similarity 20.18; Pred. No. 7.4e-05;
Matches 76; Conservative 144; Mismatches 158; Indels 0; Gaps 0;

[illegible]

```
QY 1752 GTGTCCTGGGCCACCCG 1769
      : | :::::
Db 801 SGCMCAMSSASSASSSS 818
```

LOCUS	DEFINITION	GENE	FEATURES	REFERENCE
CNS0091P	Drosophila melanogaster genome survey sequence TENG end of BAC # BACR19D16 of RPC1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	925 bp	DNA	linear GSS 03-JUN-1999

accession AL053013 GI:4934461
 version AL053013.1
 keywords GSS.
 source Drosophila melanogaster.
 organism Drosophila melanogaster.
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Phylorhodes; Drosophilidae; Drosophila.
 1 (bases 1 to 945)

AUTHORS
TITLE
JOURNAL

COMMENT

BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genome.cns.fr
- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazuhiro Oseegawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2: cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1. .925

```
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_lib="RPCI-98"
```


BASE COUNT	120 a	61 c	61 g	172 t	511 others
ORIGIN	/note="end : TERT3"				
Query Match	3.7% ; Score 67.2; DB 17; Length 925;				
Best Local Similarity	13.9% ; Pred. No. 7.3e-05;				
Matches	43; Conservative 158; Mismatches 109; Indels 0; Gaps 0;				
QY	1497 CCAGCAGTACGTGGGTGATCCCGGGGTGCTGATAGTGGCTTGATGAGTCCCGGGGTGGCCGGG 1556				
DB	601 C5SSBSBSKCSSTSBSCSCCCKSVCTGSCSSSSSSSSSSSSSSSSSSSTSSSTSSSTSSKSSSG 660				
QY	1557 CACACCTTGCTGTTCTGTATACGACACAAGTGAAGTGAAGCTCCGGCGGTGAGCAGACAGACAG 1616				
DB	661 SSSSSSSSTTTKSTKSTKSTASASSGSMAGCGSGSTGTSTSSSSSSSTSSSTSSSVSGKSSSTB 720				
QY	1617 GTGCCCGGGCGAGTCTCGAGTCGACGTCACGCCCCGACGCTCGGCCGCCCGGCTGCTCAGGCGCA 1676				
DB	721 SSGSSSSSSSSSSSSSTSSSBSCSTSSSSSSSSSVSSSSSSSCCTCCCSYSSSTSSSSSSSTS 780				
QY	1677 GCACGGGGCCCCGACGCGCTCGCGCCGACCCGCCGCCCTCCCTCCCAACCCCACTCTGGG 1736				
DB	781 WGSTGSSSSSVGTSSSSSDSTSTCCSCCCYCMCTCTYBMBEYTSCTCGSSSSSGKGGV 840				
QY	1737 CGCGGTTCCGGGGGGCGTGTCTGTGGCGACCCCGGCTTATATACGGGCGCGCGCGG 1796				
DB	841 TTCGCGCGCGSSSTNGMBGTSSACSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSVSSSSSVSSSS 900				
QY	1797 GGCGGCCCCAG 1806				
DB	901 SSASKSSSG 910				
RESULT 13	BOJ18616 1016 bp mRNA linear EST 20-AUG-2002				
LOCUS	BOJ18616				
DEFINITION	AGENCOURT 8885548 NCI CGAP CO24 Mus musculus cDNA clone				
VERSION	IMAGE:6397594 5', mRNA sequence.				
KEYWORDS	BOJ18616				
SOURCE	BOJ18616.1 GI:22333314				
ORGANISM	house mouse.				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	1 (bases 1 to 1016)				
TITLE	NIH-MGC http://mgc.nci.nih.gov/.				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)				
COMMENT	Unpublished (1999)				
	Contact: Robert Strusberg, Ph.D.				
	Email: cga@b-remail.nih.gov				
	Tissue Procurement: The Cepko Laboratory				
	cDNA Library Preparation: Life Technologies, Inc.				
	DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)				
	Clone distribution: MGC clone distribution information can be				
	found through the I.M.A.G.E. Consortium/LLNL at:				
	http://image.llnl.gov				
	Plate: L1AM13895 row: 1 column: 11				
	High quality sequence stop: 407.				
FEATURES	Location/Qualifiers				
SOURCE	1..1016				
	/organism="Mus musculus"				
	/strain="FVB/N"				
	/db_xref="taxon:10090"				
	/clone="IMAGE:6397594"				
	/clone_lib="NCI CGAP CO24"				
	/lab_host="DH10B (TI phage-resistant)"				
	/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: oligo dt.				
	Average insert size 1.6 kb. Constructed by Life				
	Technologies. Note: this is a NCI CGAP library."				
BASE COUNT	125 a	292 c	428 g	70 t	101 others

	ORIGIN	
	Query Match	3.7%; Score 67; DB 14; Length 1016;
	Best Local Similarity	50.0%; Pred. No. 8.1e-05;
	Matches 183; Conservative	0; Mismatches 182; Indels 1; Gaps 1
Oy	1442 GAGGACGGGAAGGGGTAAAGCGCGCGGCCTCAGAGATCGGGGGTGGTGAATTGGCCAGG	1501
Db	507 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGNNNGSGGGGGGGGGGGGGGGGGGGGG	566
Oy	1502 AGTGACTGGGGGTGACCGGGGGGTGCTAGGTGGCTCGAGATGCCCGGGGTGGCCGGCAC	1561
Db	567 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGNNGSGGGGGGGGGGGGGGGGGGGGG	626
Oy	1552 CTTGGTCTCTTAGACGACAAGTGACGGGGCTTCGGGGCTGACGACGAGCAAGTGCC	1621
Db	627 CGGGCGCGCGCCGCCGG	686
Oy	1622 CGGCG-CGAGTCTCGAGTGTGACCGCCCCCGCAGGTGCGGCGCGGGCTGTACGGGGC	1680
Db	667 CGGGCGCGCGGG	746
Oy	1681 GGGCGCGCGACGCGGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1740
Db	747 GCG	806
Oy	1741 CGTTCCGGGGCGGTGTCTGTGGGGCCACCCTCGCTTTAATAACGGGGCGGGCGGGCG	1800
Db	807 GCCCGCGGGCG	866
Oy	1801 GCCCAG 1806	
Db	867 CGCCCG 872	
RESULT 14		
LOCUS	BQ051092	1162 bp mRNA linear EST 29-MAR-2007
DEFINITION	AEENCOURT_6941180 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5785906	
ACCESSION	BQ051092	5' mRNA sequence.
VERSION	BQ051092.1	GI:19810432
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
JOURNAL	1 (bases 1 to 1162)	
COMMENT	NIH-MGC http://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabbs-r@mail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Plate: LLM12873 row: m column: 11 High quality sequence stop: 397. Location/Qualifiers 1..1162 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:5785906" /clone_11b="NIH_MGC_71" /tissue_type="leiomyosarcoma" /lab_host="DH10B (phage-resistant)" /note="Organ: uterus; Vector: pCMV-SPORT6; Site: 1; Note; Site: 2; Salt: Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb."	
FEATURES		
SOURCE		

BASE COUNT 71 a 373 c 538 g 102 t 78 others
 ORIGIN

Query Match 3.7%; Score 66.8; DB 14; Length 1162;
 Best Local Similarity 49.2%; Pred. No. 8.9e-05;
 Matches 178; Conservative 0; Mismatches 183; Indels 1; Gaps 1;

QY 1442 GAGAGCGGAAGGGGTAAGGCGCGGCTCAAGAGATCGGGGCTGAGTGGCCAG 1501
 Db 473 GGGGGGGGGGNGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 532
 QY 1502 AGGAGCTGGGGGTGACCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1561
 Db 533 GGGGGGGGGGNGGG 552
 QY 1562 CTGGTCTTGAGACGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1621
 Db 593 GGG 652
 QY 1622 CGGCGGAGTCTGAGCTGACGCGCGCGGAGCTGCGCGCGGCTGAGGCGGAGCAG 1681
 Db 653 GGGCGGG 712
 QY 1682 GGGCGCGGACGCTGCGCTGGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1741
 Db 713 GGGCGCGGCG 771
 QY 1742 GTTCCGGGGGCTGCTGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1801
 Db 772 GCG 831
 QY 1802 CC 1803
 Db 832 CC 833

RESULT 15
 LOCUS B0943816 1136 bp mRNA linear EST 21-AUG-2002
 DEFINITION AGENCOURT 8773466 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6372516
 5', mRNA sequence.
 ACCESSION B0943816
 VERSION B0943816.1 GI:22359294
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1136)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: DCTD/DTF/Gazdar
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LCM2548 row: g column: 13
 High quality sequence start: 54
 High quality sequence stop: 323.
 Location/Qualifiers
 1..1136
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="6372516"
 /clone_1ib="NIH_MGC_18"
 /tissue_type="large cell carcinoma"
 /lab_host="PH10B (phage-resistant)"
 /note="Organ: Lung; Vector: pOTB; Site_1: XhoI; Site_2:

ECORI: cDNA made by oligo-dT priming. Directionally cloned
 into EcorI/XhoI sites using the following 5' adaptor:
 GGCACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH-MGC Library.
 BASE COUNT 80 a 268 c 669 g 65 t 54 others
 ORIGIN

Query Match 3.7%; Score 66.6; DB 14; Length 1136;
 Best Local Similarity 47.1%; Pred. No. 9.9e-05;
 Matches 171; Conservative 0; Mismatches 192; Indels 0; Gaps 0;

QY 1442 GAGAGCGGAAGGGGTAAGGCGCGGCTCAAGAGATCGGGGCTGAGTGGCCAG 1501
 Db 544 GGG 603
 QY 1502 AGTGAAGTGGGGGTGACCGGGGGGTGAGTGGCGCTGAGGAGGAGGAGGAGGAG 1561
 Db 604 GGG 663
 QY 1562 CTGGTCTTGAGACGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1621
 Db 664 GGG 723
 QY 1622 CGGCGGAGTCTGAGCTGACGCGCGCGGAGCTGCGCGCGGCTGAGGCGGAGCAG 1681
 Db 724 GGG 783
 QY 1682 GGGCGGAGCGTGCCTGCGGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1741
 Db 784 GCG 843
 QY 1742 GTTCCGGGGGCTGCTGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1801
 Db 844 GGG 903
 QY 1802 CCC 1804
 Db 904 GCC 906

Search completed: February 23, 2003, 10:31:25
 Job time : 2400 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 23, 2003, 08:25:13 ; Search time 91 Seconds

(without alignments)
6086.354 Million cell updates/sec

Title: US-09-865-879-1

Perfect score: 1806
Sequence: 1 ggggattcgtttgtttcct.....ggcgccgcggcgcccaag 1806

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/PCRTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78.6	4.4	7218	1	US-08-232-463-14
2	56.2	3.1	329	1	US-08-253-155A-23
3	54.4	3.0	2397	1	US-07-891-9426-11
4	54	3.0	45546	4	US-09-146-053-6
5	53.2	2.9	305	1	US-08-253-155A-4
6	53	2.9	349	1	US-08-253-155A-14
7	51.6	2.9	6453	3	US-08-306-691B-14
8	51.6	2.9	6453	3	US-09-209-668-10
9	51.6	2.9	6453	3	US-09-356-952-8
10	51.2	2.8	2791	4	US-09-570-367C-1
11	50.4	2.8	10409	3	US-08-772-440-33
12	50.2	2.8	1377	2	US-08-810-572A-1
13	50.2	2.8	1377	2	US-09-290-333-1
14	50	2.8	53526	3	US-08-658-136-2
15	50	2.8	53577	3	US-08-658-136-1
16	49.4	2.7	289	4	US-09-007-005-17
17	49.4	2.7	289	4	US-09-244-796-17
18	49.4	2.7	3892	2	US-08-555-723B-3
19	49.4	2.7	3892	2	US-09-123-465-3
20	49.2	2.7	14060	4	US-08-658-136-4
21	49.2	2.7	14148	4	US-09-052-468-7
22	49	2.7	152331	3	US-09-128-155-16
23	49	2.7	176373	3	US-09-128-155-17
24	48.8	2.7	243	1	US-07-922-723A-9
25	48.8	2.7	243	1	US-07-799-828C-9
26	48.8	2.7	243	1	US-08-074-275-9
27	48.8	2.7	243	1	US-08-480-366-9

28	48.8	2.7	243	2	US-07-952-277A-9	Sequence 9, Appl1
29	47	2.6	13842	4	US-09-105-537-30	Sequence 30, Appl1
30	47	2.6	36778	4	US-09-105-537-5	Sequence 5, Appl1
31	47	2.6	38506	3	US-09-320-878-19	Sequence 19, Appl1
32	46.6	2.6	2191	1	US-08-417-330A-11	Sequence 11, Appl1
33	46.6	2.6	2440	4	US-08-724-984A-1	Sequence 1, Appl1
34	46.6	2.6	12912	2	US-08-460-751-1	Sequence 1, Appl1
35	46.2	2.6	1559	4	US-09-019-095A-7	Sequence 7, Appl1
36	46.2	2.6	2917	2	US-08-437-607A-3	Sequence 3, Appl1
37	46.2	2.6	6405	4	US-09-281-481A-18	Sequence 18, Appl1
38	45.4	2.5	291	1	US-07-922-723A-7	Sequence 7, Appl1
39	45.4	2.5	291	1	US-07-799-828C-7	Sequence 7, Appl1
40	45.4	2.5	291	1	US-08-074-275-7	Sequence 7, Appl1
41	45.4	2.5	291	1	US-08-480-366-7	Sequence 7, Appl1
42	45.4	2.5	291	2	US-07-952-277A-7	Sequence 7, Appl1
43	45.4	2.5	2280	1	US-08-415-818-5	Sequence 5, Appl1
44	45.4	2.5	2280	2	US-08-894-236-5	Sequence 5, Appl1
45	45.4	2.5	2280	5	PCT-US96-01444-5	Sequence 5, Appl1

ALIGNMENTS

RESULT 1
US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
TITLE OF INVENTION: RECOMBINANT FOMLPD VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMUG
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-F1s
US-08-232-463-14
Query Match 4.4%; Score 78.6; DB 1; Length 7218;

	Best Local Matches	Similarity	11.2%; Conservative	Pred. No. 159;	3.6e-11; Mismatches	Indels	Gaps	
OY	256	TGTAGAGTCCATTATTTATTTACTACAAAGGAGAGAATAATTATCTTCANT						315
Db	1486	TGTAGCACTACCTGTAAATTACCTATCTAAGCAAGTGTTAAAGATGAAGAAATTGGT						1437
OY	316	CTAAAAGGAGCATATATATGTATTAAGAGAGCGTATATATATGGAGAGAGTGAGC	:	:	:	:	:	375
Db	1436	ACRRR	:	:	:	:	:	1377
OY	376	AGGGCGAAAAAGGAGAGAGCAGACAGAGAAAGGAGGAGAGACAGAGACAC	:	:	:	:	:	435
Db	1376	RRR	:	:	:	:	:	1317
OY	436	TGGCGAGAGATCGATTAGAGAGACAGAAATGATGAATGAAGATTAACTTCACCACAAG	:	:	:	:	:	495
Db	1316	RRR	:	:	:	:	:	1257
OY	496	CTTGCCTGAGAGGGAATGAGAGAG	:	:	:	:	:	522
Db	1256	RRR	:	:	:	:	:	1230

```

: RESULT 2
: US-08-253-155A-23
: Sequence 23, Application US/08253155A
: Patent No. 5691147
:
: GENERAL INFORMATION:
: APPLICANT: Gyuris, Jeno
: APPLICANT: Draetta, Giulio
: TITLE OF INVENTION: CDK4 Binding Proteins
: NUMBER OF SEQUENCES: 95
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: LAHIVE & COCKFIELD
: STREET: 60 State Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: ASCII(text)
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/253,155A
: FILING DATE: 02-JUN-1994
: CLASSIFICATION: 435
:
: ATTORNEY/AGENT INFORMATION:
: NAME: Vincent, Matthew P.
: REGISTRATION NUMBER: 36,709
: REFERENCE/DOCKET NUMBER: MII-028
:
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 227-7400
: TELEFAX: (617) 227-5941
:
: INFORMATION FOR SEQ ID NO: 23:
:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 329 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: MOLECULE TYPE: CDNA
:
: US-08-253-155A-23

```

	Query Match	3.1%	Score 56.2	DB 1	length 329
	Best Local Similarity	61.1%	Pred. No. 8.7e-06		
	Matches	91	Conservative	0	Mismatches 95
					Indels 0
					Gaps 0
Qy	340	AAAGCAACTGTTATTCATTCATTCGGGGCAAGAGCTGGACAGGGGGAAAAAGGCGACAGACGACG	399		
Db	25	AGAGGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG	84		

OY	ACAGAAATGATGAATGCAGATTAACTTCA	488
OY		
Db	ACAGAGAGTCCTCATGATCTTTCCACTTCA	173

```

RESULT 3
US-07-891-942G-11
: Sequence 11, Application US/07891942G
: Patent No. 5679511
: GENERAL INFORMATION:
: APPLICANT: Kwon, Byoung Se
: TITLE OF INVENTION: CDNA CLONES FOR HUMAN TYROSINASE AND FOR
: TITLE OF INVENTION: A REGULATORY PROTEIN IN THE MELANIN PROTEIN PATHWAY
: NUMBER OF SEQUENCES: 16
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Christopher A. Michaels, Barnard, Brown &
: ADDRESSEE: Michaels
: STREET: 306 East State Street; Suite 220
: CITY: Ithaca
: STATE: NY
: COUNTRY: USA
: ZIP: 14850
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/891,942G
: FILING DATE: 01-JUN-1992
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 06/915,753
: FILING DATE: 06-OCT-1986
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/362,847
: FILING DATE: 07-JUN-1989
: ATTORNEY/AGENT INFORMATION:
: NAME: Michaels, Christopher A
: REGISTRATION NUMBER: 34,390
: REFERENCE/DOCKET NUMBER: INDI
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 607-273-1711
: TELEFAX: 607-273-2609
: INFORMATION FOR SEQ ID NO: 11:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2397 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: US-07-891-942G-11

```

	Query Match	3.0%;	Score 54.4;	DB 1;	Length 2397;	
	Best Local Similarity	61.1%;	Pred No. 6.9e-05;			
	Matches	88;	Conservative	0;	Mismatches	56; Indels 0; Gaps 0
QY	319	AAAGGACACTATATATGTATATAAAGCAACCTTATATGTTGGGGAAGAAGTGTCAGC	378			
Dd	1466	AGAGGAGAGAGAGAGAGAGATAGAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGACA	1525			
QY	379	GCGAAAAGGGGAGAGACGAGAGACGAAAAAGGAGGAGAGGCACAAAGAGACACACTGG	438			
Dd	1526	GCGGAGACAGAGAGAGAGAGAGGAGGAGACAGACAGAACAAAGACAGAGAGAGGAGAGACA	1585			
QY	439	GCGAGAGATCGATTAGAGAGACA	462			
Dd	1586	GAGAGAGACCTTTTAACGTGAGATA	1609			

RESULT 4
US-09-146-053-6
Sequence 6, Application US/09146053A
Patent No. 6399349
GENERAL INFORMATION:
APPLICANT: Ryan, James W.
APPLICANT: Sprinkle, Terry Joe Curtis
APPLICANT: Venema, Richard C.
TITLE OF INVENTION: Human Aminopeptidase P Gene
FILE REFERENCE: MCG103
CURRENT APPLICATION NUMBER: US/09/146, 053A
CURRENT FILING DATE: 1998-09-02
EARLIER APPLICATION NUMBER: 60/057, 854
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 45546
TYPE: DNA
ORGANISM: Homo sapiens
US-09-146-053-6

Query Match 3.0%; Score 54; DB 4; Length 45546;
Best Local Similarity 56.0%; Pred. No. 0.00037;
Matches 102; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 318 AAAGAGAGTATATATGTAAGAGAGCTGTATGATATGAGGAGAGGTGACAG 377
DB 16595 AGAAG 18654
QY 378 GGGGAG 437
DB 18655 AG 18714
QY 438 GCGGAGAGATCGATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 497
DB 18715 AGAAG 18774
QY 498 TC 499
DB 18775 TC 18776

RESULT 5
US-08-253-155A-4
Sequence 4, Application US/08253155A
Patent No. 5691147
GENERAL INFORMATION:
APPLICANT: Gyuris, Jeno
APPLICANT: Draetta, Giulio
TITLE OF INVENTION: CDK4 Binding Proteins
NUMBER OF SEQUENCES: 95
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/253, 155A
FILING DATE: 02-JUN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MIT-028

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 305 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-253-155A-4

Query Match 2.9%; Score 53.2; DB 1; Length 305;
Best Local Similarity 63.1%; Pred. No. 5.4e-05;
Matches 82; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 352 ATAGATATGGGGAGAGAGGTGACAGGGGAGAGAGAGAGAGAGAGAGAGAG 411
DB 21 AG 80
QY 412 AGGAG 471
DB 81 AG 140
QY 472 AATGAGATT 481
DB 141 AGAGAGATT 150

RESULT 6
US-08-253-155A-14
Sequence 14, Application US/08253155A
Patent No. 5691147
GENERAL INFORMATION:
APPLICANT: Gyuris, Jeno
APPLICANT: Draetta, Giulio
TITLE OF INVENTION: CDK4 Binding Proteins
NUMBER OF SEQUENCES: 95
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/253, 155A
FILING DATE: 02-JUN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MIT-028
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 349 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-253-155A-14

Query Match 2.9%; Score 53; DB 1; Length 349;
Best Local Similarity 58.1%; Pred. No. 6.5e-05;
Matches 93; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 340 AAGGAAGCTGTATAGATATGCGGAGAGAGTGCACAGCGGGAAGAGAGAGAGAG 399
DB 51 AG 110
QY 400 AG 459
DB 111 AG 170
QY 460 ACAGAAATGATGATGAATTAACCTTCACCCAGGCTTC 499
DB 171 AGAGAGCCAGGCTTAAACATATGAGACTGATGTATC 210

RESULT 7
US-08-306-691B-14
; Sequence 14, Application US/08306691B
; Patent No. 5734039
; GENERAL INFORMATION:
; APPLICANT: Calabretta, Bruno
; APPLICANT: Skorski, Tomasz
; TITLE OF INVENTION: ANTISENSE
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavoragna & Monaco, P.C.
; STREET: Two Penn Center, Suite 1800
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/306,691B
; FILING DATE: September 15, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5734039e
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6453 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-306-691B-14

Query Match 2.9%; Score 51.6; DB 1; Length 6453;
Best Local Similarity 46.2%; Pred. No. 0.00063;
Matches 171; Conservative 0; Mismatches 199; Indels 0; Gaps 0;

QY 1406 CCGGACACCGCTCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1465
DB 611 CCGCTCGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 670
QY 1466 GGGGTGAAGAGATCGGGGGTGTGAGTGGCCAGAGAGTACGTGGGTGACCGGGGGTGC 1525
DB 671 TGAGGTGAAG 730
QY 1526 TGAGGTGGCTGAGATGCGGGGTGCGGGGAGACACTTGTGTTTGTAGACGACAAAGT 1585

DB 731 TGAGCGGCGCTCGGGGCTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 790
QY 1586 GACGGGCTCGGGGCTGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1645
DB 791 TCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 850
QY 1646 CCGGAGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1705
DB 851 CTCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCA 910
QY 1706 ACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1765
DB 911 CGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 970
QY 1766 CCGGCGGCTTC 1775
DB 971 AGTGGCTTT 980

RESULT 8
US-09-209-668-10
; Sequence 10, Application US/09209668A
; Patent No. 6114517
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Xu, Xiaoxing S.
; TITLE OF INVENTION: METHODS OF MODULATING TUMOR NECROSIS FACTOR
; FILE REFERENCE: ISPH-0336
; CURRENT APPLICATION NUMBER: US/09/209,668A
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 6453
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1664)..(1774)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2042)..(2220)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2374)..(2533)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3231)..(3350)
; DATABASE ACCESSION INFORMATION:
; DATABASE ENTRY DATE: 1991-01-03
; US-09-209-668-10

Query Match 2.9%; Score 51.6; DB 3; Length 6453;
Best Local Similarity 46.2%; Pred. No. 0.00063;
Matches 171; Conservative 0; Mismatches 199; Indels 0; Gaps 0;

QY 1406 CCGGACACCGGCTCGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1465
DB 611 CCGCTCGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 670
QY 1466 GGGGTGAAGAGATCGGGGGTGTGAGTGGCCAGAGAGTACGTGGGTGACCGGGGGTGC 1525
DB 671 GGGGTGAAG 730
QY 1526 TGAGGTGGCTGAGATCGGGGGTGTGAGTGGCCAGAGAGTACGTGGGTGACCGGGGGTGC 1585
DB 731 TGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 790
QY 1586 GACGGGCTCGGGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1645
DB 791 TCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 850

```

QY 1646 CCGAGCTCGGCCCCGCTCTCTAGGCGGAGACAGCGGCCCCCGACCGCTGCTGCGCG 1705
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 851 CTCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 910
QY 1706 ACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1765
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 911 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 970
QY 1766 CCGGCGCTCT 1775
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 971 AGTGCGCTTTT 980

```

```

RESULT 9
US-09-356-952-8
; Sequence 8, Application US/09356952
; Patent No. 6117663
; GENERAL INFORMATION:
; APPLICANT: Borlack-Stodin, Ann
; APPLICANT: Margalit, S. M.
; APPLICANT: Bor-Sogil, Dafna
; APPLICANT: Cole, Phillip
; APPLICANT: Kurtyan, John
; TITLE OF INVENTION: A CRYSTAL OF A RAS-SOS COMPLEX AND METHODS OF USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 600-1-228N
; CURRENT APPLICATION NUMBER: US/09/356,952
; EARLIER FILING DATE: 1999-07-19
; EARLIER APPLICATION NUMBER: 60/099,631
; EARLIER FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 8
; LENGTH: 6453
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-356-952-8

```

```

Query Match
Best Local Similarity 46.2%; Score 51.6; DB 3; Length 6453;
Matches 171; Conservative 0; Mismatches 199; Indels 0; Gaps 0;

```

```

QY 1406 CCCCCGACCGGCTCGCGCGCGGAGAGACTCACCCGAGAGCGGAAGGGTAAAGGCGCG 1465
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 611 CTTCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 670
QY 1466 GGGGTGAAGAGATCGGGGGTGTGAGTTGGCCAGAGTACTGGGGTGAACCGGGGTGC 1525
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 671 GGGGTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 730
QY 1526 TGAGGTGCTGAGTGCAGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1585
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 731 TGAGCGGGGCGTGGGGGCGTGGGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGGAA 790
QY 1586 GAGCGGCTCGGGCGTGAAGCAGAGAGAGAGTCCCGGGCGAGTCTGAAGTGCACAGCC 1645
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 791 TCCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 850
QY 1646 CCGGAGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1705
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 851 CTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 910
QY 1706 ACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1765
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 911 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 970
QY 1766 CCGGCGCTCT 1775
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 971 AGTGCGCTTTT 980

```

RESULT 10

```

US-09-570-367C-1/c
; Sequence 1, Application US/09570367C
; Patent No. 6338851
; GENERAL INFORMATION:
; APPLICANT: Gorczynski, Reginald M.
; TITLE OF INVENTION: Methods and Compositions for Immunomodulation
; FILE REFERENCE: 9579-21
; CURRENT APPLICATION NUMBER: US/09/570,367C
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/064,764
; PRIOR FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 1
; LENGTH: 2791
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-570-367C-1

```

```

Query Match
Best Local Similarity 57.5%; Score 51.2; DB 4; Length 2791;
Matches 92; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

```

```

QY 318 AAAAGCAGTATATATATATATATATATATATATATATATATATATATATATATATAT 377
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 382 AGACAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 323
QY 378 GGGGAAAGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 437
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 322 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 263
QY 438 GCGGAGATCGATTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 262 AAAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 223

```

```

RESULT 11
US-08-772-440-33
; Sequence 33, Application US/08772440
; Patent No. 6046158
; GENERAL INFORMATION:
; APPLICANT: Aritizumi, Kiyoshi
; APPLICANT: Takashima, Akira
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
; TITLE OF INVENTION: LECTINS, DCTIN-1 AND DCTIN-2; COMPOSITIONS AND USES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/772,440
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTXD:493
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/474-7577
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10409 base pairs

```

```

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: modified_base
LOCATION: 6510
OTHER INFORMATION: /mod_base- OTHER
OTHER INFORMATION: /note- "D - A or G or T"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3406..6470
OTHER INFORMATION: /mod_base- OTHER
OTHER INFORMATION: /note- "K - G or T"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3564..7896
OTHER INFORMATION: /mod_base- OTHER
OTHER INFORMATION: /note- "M - A or C"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3497..3607
OTHER INFORMATION: /mod_base- OTHER
OTHER INFORMATION: /note- "N - A or C or G or T"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3479..6422
OTHER INFORMATION: /mod_base- OTHER
OTHER INFORMATION: /note- "R - A or G"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3405..6871
OTHER INFORMATION: /mod_base- OTHER
OTHER INFORMATION: /note- "S - C or G"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3457..9998
OTHER INFORMATION: /mod_base- OTHER
OTHER INFORMATION: /note- "W - A or T"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3595..9999
OTHER INFORMATION: /mod_base- OTHER
OTHER INFORMATION: /note- "Y - C or T"
US-08-772-440-33

```

```

Query Match 2.8%; Score 50.4; DB 3; Length 10409;
Best Local Similarity 50.4%; Pred. No. 0.0017;
Matches 123; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 226 CATGTTCTGCTTGGCTCTTTTATTTAGAGTCCATTTATTTATTTAGT 285
DB 4904 CTTGTTCTCTGTCAGTTCAGTCTGCTGCTTCTTCTTTCTTTAAATTTT 4963

QY 286 ACAAGGGGAGCAAAATGATCTTTCAATCAAAAGGAGCTATATATGATAAAGGA 345
DB 4964 AAATTTGTCTTTAAATAATTTGCCCCGTTATTTTCAAGAGAGAGAGAGAGAGAG 5023

QY 346 AGCTGATATGATATGGGAGAGAGTGGACAGGGGAAAAGGGGAGAGAGAGAGAG 405
DB 5024 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5083

QY 406 AAAGGAGGAGAGGAGCAAGAGAGAGAGAGAGAGAGAGAGATCGATTAGAGAGAGAG 465
DB 5084 AGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTA 5143

QY 466 ATGA 469
DB 5144 TGA 5147

```

RESULT 12
US-08-810-572A-1
; Sequence 1, Application US/08810572A

```

Patent No. 5969102
GENERAL INFORMATION:
APPLICANT: Bram, Richard J.
APPLICANT: von Bulow, Gotz
TITLE OF INVENTION: A LYMPHOCYTE SURFACE RECEPTOR THAT BINDS
TITLE OF INVENTION: CAML, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF USE
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
STREET: floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/810,572A
FILING DATE: 28-FEB-1997
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1340-1-007
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1377 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-810-572A-1

```

```

Query Match 2.8%; Score 50.2; DB 2; Length 1377;
Best Local Similarity 62.2%; Pred. No. 0.00071;
Matches 79; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 346 AGCTGATATGATATGGGAGAGAGTGGACAGGGGAAAAGGGGAGAGAGAGAG 405
DB 1082 AGAGGAAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1141

QY 406 AAAGGAGGAGAGGAGCAAGAGAGAGAGAGAGAGAGAGAGATCGATTAGAGAGAGAG 465
DB 1142 AGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1201

QY 466 ATGATGA 472
DB 1202 AGCAGGA 1208

```

RESULT 13
US-09-333-1
; Sequence 1, Application US/09290333
; Patent No. 6316222
GENERAL INFORMATION:
APPLICANT: Bram, Richard J.
APPLICANT: von Bulow, Gotz
TITLE OF INVENTION: A LYMPHOCYTE SURFACE RECEPTOR THAT BINDS
CAML, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF USE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:

ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/290,333
FILING DATE: 12-Apr-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1340-1-007 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1377 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-290-333-1

Query Match 2.8%; Score 50.2; DB 4; Length 1377;
Best Local Similarity 62.2%; Pred. No. 0.00071;
Matches 79; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 346 AGCTGTATAGATATGAGGAGAGAGTGGACAGAGGAGAAAGGAGAGAGAGAGAG 405
DB 1082 AGAGGAGAAAGCAGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1141
QY 406 AAAGGAGGAG 465
DB 1142 AGAGGAG 1201
QY 466 ATGATGA 472
DB 1202 AGCAGGA 1208

RESULT 14
US-08-658-136-2/c
Sequence 2, Application US/08658136
Patent No. 6071717
GENERAL INFORMATION:
APPLICANT: KLINGER, KATHERINE W
APPLICANT: LANDES, GREGORY M
APPLICANT: BURN, TIMOTHY C
APPLICANT: CONNORS, TIMOTHY D
APPLICANT: DACKOWSKI, WILLIAM
APPLICANT: GERMINO, GREGORY
APPLICANT: QIAN, FENG
TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENZYME CORPORATION
STREET: ONE MOUNTAIN ROAD
CITY: FRAMINGHAM
STATE: MASSACHUSETTS
COUNTRY: USA

ZIP: 01701

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/658,136
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: LASSEN, ELIZABETH
REGISTRATION NUMBER: 31,845
REFERENCE/DOCKET NUMBER: GEN4-17.8
TELECOMMUNICATION INFORMATION:
TELEPHONE: 508-872-8400
TELEFAX: 508-872-5415
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 53526 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

US-08-658-136-2

Query Match 2.8%; Score 50; DB 3; Length 53526;
Best Local Similarity 49.2%; Pred. No. 0.0047;
Matches 131; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

QY 1538 GAGTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1597
DB 3887 GGGGCGCCGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3828
QY 1598 GCGTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1657
DB 3827 ACCGAGGCTCCGAGCCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3768
QY 1658 CCGGCTGCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1717
DB 3767 GCGGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3708
QY 1718 CCGAAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1777
DB 3707 CCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1748
QY 1778 ATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1803
DB 3647 CGTTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3622

RESULT 15
US-08-658-136-1/c
Sequence 1, Application US/08658136
Patent No. 6071717
GENERAL INFORMATION:
APPLICANT: KLINGER, KATHERINE W
APPLICANT: LANDES, GREGORY M
APPLICANT: BURN, TIMOTHY C
APPLICANT: CONNORS, TIMOTHY D
APPLICANT: DACKOWSKI, WILLIAM
APPLICANT: GERMINO, GREGORY
APPLICANT: QIAN, FENG
TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENZYME CORPORATION
STREET: ONE MOUNTAIN ROAD
CITY: FRAMINGHAM
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 01701
COMPUTER READABLE FORM:

```

;
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/658,136
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: LASSEN, ELIZABETH
; REGISTRATION NUMBER: 31,845
; REFERENCE/DOCKET NUMBER: GEN4-17.8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 508-872-8400
; TELEFAX: 508-872-5415
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 53577 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-658-136-1

```

```

Query Match      2.88; Score 50; DB 3; Length 53577;
Best Local Similarity 49.2%; Pred. No. 0.0047;
Matches 131; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

QY 1538 GAGTGGCGGGGGTGGCGGACACCTTGTTCTGTAGACGACAAGTGAGCGGCTCCGG 1597
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 3887 GGGTGGCGGCTGGGGCCCTACTACACAGCGCTGTGCTCCCGGGGATGCCAGCGGGG 3828
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 1598 GCGTGAGCAGAGAGAGAGTGCCTGGGCGAGTCTGAGCTGACAGGCCCGGAGCTCGGC 1657
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 3827 ACCGAGCGTCCGAGCGCGCGGCGCGAGCAATTGACGGCGGAGCGGGCGGCGCTGG 3768
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 1658 CCGGCGTCTCAGGGCGAAGCAGGGGCCCGGACGCTGCTGCGCCGACCGCCCT 1717
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 3767 GCGCGAGAGCAGGGGGGCTCGCAGGGCCCGACGCGCGCGGGGGCGCCCGCCAGCGC 3708
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 1718 CCCAACCCCTCTCTGGGCGCGGTCTCGGGGCGTGTCTTGGGCAACCCCGGCTTCTAT 1777
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 3707 CCGAGCCACAGGCGCCAGGCGCCAGGCGCCAGCGCGGGCGGGCGGGCGGCGCAT 3648
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 1778 ATACGGGCGCGCGCGCGCGCGCC 1803
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 3647 CGTAGGCGAGCGCGCGCATGGCCCC 3622
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: February 23, 2003, 10:37:08
 Job time : 353 secs

Qy	1	GGGATTCGTTTTGTTTTCTTCATTTTCCAAATAAATCAGAGATCCGTCTTGAGGTG	60
Db	101	GGGGATTCGTTTTGTTTTCTTCATTTTCCAAATAAATCAGAGATCCGTCTTGAGGTG	166
Qy	61	CAAGCGAGATCTGTAAGAGAGAGTATACAGAAAGG-ATACGCAAGCGACATTATGG	119
Db	161	CAAGCGAGATCTGTAAGAGAGTATACAGAAAGAGAAACACCAAGCCAGCTATATGG	220
Qy	120	CACGGTTTCTGTAAACAGGTTGAGTGTAGCCACAGCCTGAGCACTGTGTGGGAAAGAGC	179
Db	221	CACGGTTTCTGTAAACAGGTTGAGTGTAGCCACAGCCTGAGCAGCTGTGTGGGAAAGAGC	280


```
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (390986)..(391005)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (346860)..(346823)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (317174)..(317193)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (280353)..(280373)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (271829)..(271848)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (183872)..(183891)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (170625)..(170645)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (132680)..(132700)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
OTHER INFORMATION: n is a, c, g, or t
US-09-771-208-20
```

```
Query Match 3.6%; Score 65.6; DB 9; Length 659158;
Best Local Similarity 63.18; Pred. No. 3.7e-06;
Matches 101; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
```

```
QY 318 AAAGAGAGTATATATATATAAAGAGCTGTAGATATGAGGGAAGGTGGACAG 377
||||| ||| ||||| | ||| ||| ||| ||| |||
DB 463003 AAAAGAGAGATATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 462944
QY 378 GGGGAAAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 437
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 462943 AGGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 462884
QY 438 GCGGAGAGATCGATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 477
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 462883 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 462844
```

RESULT 3
US-09-771-208-20

```
; Sequence 20, Application US/09771208
; Patent No. US2002015564A1
; GENERAL INFORMATION:
; APPLICANT: MEDRANO, JUAN
; APPLICANT: BRADFOR, ERIC
; APPLICANT: HORVAT, SIMON
; TITLE OF INVENTION: CLONING OF A HIGH-GROWTH GENE
; FILE REFERENCE: 407-923710US
; CURRENT APPLICATION NUMBER: US/09/771,208
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 08/999,477
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 659158
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (123459)..(123478)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (602466)..(602485)
; OTHER INFORMATION: n is unidentified a, c, g, or t
```

```
NAME/KEY: misc_feature
LOCATION: (546998)..(547017)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (494715)..(494814)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (390986)..(391005)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (346860)..(346823)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (317174)..(317193)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (280353)..(280373)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (271829)..(271848)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (183872)..(183891)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (170625)..(170645)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (132680)..(132700)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
OTHER INFORMATION: n is a, c, g, or t
US-09-771-208-20
```

```
Query Match 3.2%; Score 58.6; DB 9; Length 659158;
Best Local Similarity 64.2%; Pred. No. 0.0003;
Matches 88; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
```

```
QY 340 AAAGAGAGCTGTATATATATAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 399
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 343768 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 343827
QY 400 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 459
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 343828 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 343887
QY 460 ACAGAGAGATGATGATGA 476
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 343888 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 343904
```

RESULT 4

```
US-09-938-842A-3578/c
; Sequence 3578, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tony
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SAME, AND METHODS OF USE
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 3578
```

```

; LENGTH: 1721
;
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-3578

```

Query Match	3.28;	Score 57.6;	DB 9;	Length 1721;
Best Local Similarity	58.9%;	Pred. No. 1.6e-05;		
Matches 99;	Conservative 0;	Mismatches 69;	Indels 0;	Gaps 0

[illegible]

RESULT 5
US-09-867-701-9416
1. Commerce 0416 April 2009 pg 00967701

```

? APPLICANT: Aglate, Paul A.
? APPLICANT: Jones, Robert
? APPLICANT: Harlocker, Susan L.
? TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
? OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
? FILE REFERENCE: 210121.497
? CURRENT APPLICATION NUMBER: US/09/867,701
? CURRENT FILING DATE: 2001-05-29
? NUMBER OF SEQ ID NOS: 10912
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 9416
? LENGTH: 307
? TYPE: DNA
? ORGANISM: Homo sapien
? US-09-867-701-9416

```

Query Match	3.18;	Score 56.6;	DB 10;	Length 307;
Best Local Similarity	60.98;	Pred. No. 1.1e-05;		
Matches 92; Conservative	0;	Mismatches 59;	Indels 0;	Gaps 0;

Qy	316	CTAAAGCAGCAGTATTATTTCTATATAAAAGGAAGCTGTATAGCTATATGGGGAAAGAGTGGAC	375
Db	10	CTGGCAGAAAGCAGCAGCAGGAAGCGAG	69
Qy	376	AGGGGGGAAAAGGGGAGANGAGNCAGNAGNAGNAGAAAAGGGAGGAGAGGGCAAGAGAGAGAC	435
Db	70	AGAGAGAGAGGGCCGAGACCGCAGACACACAGACAGAGAGAGACAGAGAGAGAGAGAGAGAGAGAG	129
Qy	436	TGGCGCAGACATCCATTTCAGAGACACAGAAA	466
Db	130	AGGGCGAGAGCGAGAGAGACAGCAGAGAGAGAGA	160

RESULT 6
US-09-867-701-8978
: Sequence 8978, Application US/09867701
: Patent No. US20020133237A1
: GENERAL INFORMATION
: APPLICANT: Aglate, Paul A.
: APPLICANT: Jones, Robert
: APPLICANT: Harlocke, Susan L.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: TITLE OF INVENTION: COMPOSITIONS AND DIAGNOSIS OF OVARIAN CANCER
: FILE REFERENCE: 210121.497
: CURRENT APPLICATION NUMBER: US/09/867,701
: CURRENT FILING DATE: 2001-05-29

```

: NUMBER OF SEQ ID NOS: 10912
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO: 8978
: LENGTH: 298
: TYPE: DNA
: ORGANISM: Homo sapien
US-09-867-701-8978

Query Match      3.1%  Score 55.6; DB 10; Length 298;
Best Local Similarity 63.4%; Pred. No. 2e-05;
Matches 85; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Oy 351 TATACATATGCGGGAAGAGGTGACAGGCGGCAAAAAGCGAGAGCAGAGAGCAAGG 410
    |||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 20 TATACGCGCTGCAGAGACGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGA 79
    || |||| | | | | | | | | | | | | | | | | | | | | | | | |
Oy 411 GAGGAGAGAGGCAAGAGAGAGACACTGGGCGAGAGATCGATTAGGAGAGACAAATGAT 470
    || |||| | | | | | | | | | | | | | | | | | | | | | | | |
Db 80 GAGAGAGAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA 139
    || || | | | | | | | | | | | | | | | | | | | | | | | |

Oy 471 GATGAGAGATTAC 484
    || || | | | | | |
Db 140 GAGAGAAAAAGAGC 153

RESULT 7
US-09-799-462A-17/c
: Sequence 17, Application US/09799462A
: Patent No. US20020160970A1
GENERAL INFORMATION:
APPLICANT: Hadlaczky, Gyula
Scalay, Aladar
TITLE OF INVENTION: ARTIFICIAL CHROMOSOMES, USES THEREOF
AND METHODS FOR PREPARING ARTIFICIAL CHROMOSOMES
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESS: Heller Ehtman White & McLaughlin
STREET: 4250 Executive Square, 7th Floor
CITY: LA Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/799,462A
FILING DATE: 10-Sep-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/835,682
FILING DATE: 10-APR-1997
APPLICATION NUMBER: 08/695,191
FILING DATE: 07-AUG-1996
APPLICATION NUMBER: 08/682,080
FILING DATE: 15-JUL-1996
APPLICATION NUMBER: 08/629,822
FILING DATE: 10-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24601-4026
TELECOMMUNICATION INFORMATION:
TELEPHONE: 858-450-8403
TELEFAX: 858-587-5360
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 42999 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

```

Oy	351	TATGATATATGGGCGAAGAGTGTGGACAGCGGGGAAAAAGGGAGAGACACGAGAGAGAAAGC	410
Db	20	TATAGCGCTGCGAGAAAGCAGCAAGAGGGAGAGAGAGAGAGAGACAGACAGAGAGAGAGA	79
Oy	411	GAGGAGAGGCGNCAGAGAGACACTGGGCGAGAGATTCGATTAGAGAGAGCAAAATGAT	470
Db	80	GAGAGAGAGACAGAGAGAGAGAGAGAGAGAGCGCCAGACGAGACAGACAGACAGAGAGAGA	139
Oy	471	GAATGAAAGATTAAAC	484
Db	140	GAGAGAAAAAGAGC	153

RESULT 7
US-09-799-462A-17/c
; sequence 17, Application US/09799462A

APPLICANT: Hadjaczky, Gyula
Szalay, Aladar
TITLE OF INVENTION: ARTIFICIAL CHROMOSOMES, USES THEREOF
AND METHODS FOR PREPARING ARTIFICIAL CHROMOSOMES

NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:

ADDRESS: Heller Ehrman White & McCalliff
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA

```

;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Diskette
;
; COMPUTER: IBM Compatible
;
; OPERATING SYSTEM: DOS
;
; SOFTWARE: FaastsEQ Version 1.5
;

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/799,462A
; FILING DATE: 10-Sep-2001

```

PRIOR APPLICATION DATA:

1 APPLICATION NUMBER: 08/835,682
2
3 FILING DATE: 10-APR-1997
4
5 APPLICATION NUMBER: 08/695,191
6
7 FILING DATE: 07-AUG-1996
8
9 APPLICATION NUMBER: 08/682,080
0
1 FILING DATE: 15-JUL-1996
2
3 APPLICATION NUMBER: 08/629,822
4

ATTORNEY/AGENT INFORMATION:

NAME: Seldman, Stephanie L
REGISTRATION NUMBER: 33,779

REFERENCE/DOCKET NUMBER: 24601-402G
TELECOMMUNICATION INFORMATION:

TELEPHONE: 858-450-84
TELEFAX: 858-587-5360

TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 17:
;

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 42999 base pairs

```

TYPE: nucleic acid
STRANDEDNESS: single

PRIOR FILING DATE: 2001-01-30

460 ACAGAAATGATGAA 473

STIC-Biotech/ChemLib

86812

Fr m: Yu, Misook
Sent: Friday, February 14, 2003 9:22 AM
To: STIC-Biotech/ChemLib
Subject: 09/865,879

RECEIVED

FEB 14 2003

STIC-BIOTECH/ChemLib
(STIC)

Please search SEQ ID NO:2

Examiner Misook Yu, Ph.D.
703-308-2454 (Phone)
Art Unit 1642
CM1-8E18 (Room)
CM1-8E12 (Mail Box)

POINT OF CONTACT:
PAUL SCHULWITZ
TECHNICAL INFO. SPECIALIST
CM1 6B06 TEL. (703) 305-1954

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 2/20
Date Completed: 2/24
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 23, 2003, 06:17:17 ; Search time 4104 Seconds

(without alignments)
11012.831 Million cell updates/sec

Title: US-09-865-879-2

Perfect score: 1553
Sequence: 1 gactcagggtgtcccaacc.....ggtcgcgtcgtccgcctcc 1553

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*
1: gb_da:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pln:*
35: em_hlg_rod:*
36: em_hlg_mam:*
37: em_hlg_vrl:*
38: em_sy:*
39: em_hlgo_hum:*
40: em_hlgo_mus:*
41: em_hlgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1553	100.0	1553	6	AX323410	AX323410 Sequence
2	1542	99.3	76007	9	AC0004503	AC0004503 Homo sapi
3	105	6.8	2748	9	AK094581	AK094581 Homo sapi
4	96.6	6.2	7218	6	I66494	I66494 Sequence 14
5	75.8	4.9	207420	2	AC078884	AC078884 Mus muscu
6	73	4.7	151283	9	AC107540	AC107540 Homo sapi
7	71	4.6	2583	9	AK094055	AK094055 Homo sapi
8	70.2	4.5	86057	2	AC131553	AC131553 Rattus no
9	65.4	4.2	100740	2	AC131216	AC131216 Rattus no
10	65	4.2	2659	9	AK093916	AK093916 Homo sapi
11	65	4.2	119989	2	AC119472	AC119472 Rattus no
12	62.4	4.0	132222	2	AC121192	AC121192 Rattus no
13	62.4	4.0	147622	2	AC110437	AC110437 Rattus no
14	61.8	4.0	201650	2	AC119304	AC119304 Rattus no
15	61.8	4.0	203343	2	AC131133	AC131133 Rattus no
16	61.6	4.0	194562	2	AC112394	AC112394 Rattus no
17	61.4	4.0	69051	2	AC073171	AC073171 Homo sapi
18	61.2	3.9	52683	2	AC131481	AC131481 Rattus no
19	61	3.9	131956	2	AC131361	AC131361 Rattus no
20	61	3.9	190626	2	AC098181	AC098181 Rattus no
21	61	3.9	206773	2	AC126296	AC126296 Rattus no
22	60.8	3.9	106649	2	AC122626	AC122626 Rattus no
23	60.8	3.9	196828	2	AC127709	AC127709 Rattus no
24	60.6	3.9	182269	2	AC107416	AC107416 Rattus no
25	60.6	3.9	232603	2	AC127560	AC127560 Mus muscu
26	60.2	3.9	69636	2	AC129731	AC129731 Rattus no
27	60.2	3.9	112289	2	AC125732	AC125732 Rattus no
28	60	3.9	141288	2	AC116264	AC116264 Rattus no
29	59.8	3.9	204431	2	AC096927	AC096927 Rattus no
30	59.2	3.8	173022	2	AC128405	AC128405 Rattus no
31	59.2	3.8	17801	2	AC114391	AC114391 Rattus no
32	59	3.8	2790	6	AX281844	AX281844 Sequence
33	58.6	3.8	121077	2	AC105558	AC105558 Rattus no
34	58.6	3.8	152347	2	AC021096	AC021096 Homo sapi
35	58.4	3.8	143786	2	AC112363	AC112363 Rattus no
36	58.4	3.8	197949	2	AC109737	AC109737 Rattus no
37	58.2	3.7	132884	2	AC128403	AC128403 Rattus no
38	58	3.7	150305	2	AC114705	AC114705 Rattus no
39	58	3.7	166641	2	AC126966	AC126966 Rattus no
40	58	3.7	182225	2	AC128440	AC128440 Rattus no
41	58	3.7	186906	2	AC110640	AC110640 Rattus no
42	57.8	3.7	105620	2	AC127760	AC127760 Rattus no
43	57.8	3.7	133052	2	AC112614	AC112614 Rattus no
44	57.8	3.7	140340	2	AC108584	AC108584 Rattus no
45	57.6	3.7	145592	2	AC119341	AC119341 Rattus no

ALIGNMENTS

RESULT 1
AX323410
LOCUS AX323410 1553 bp DNA linear PAT 07-JAN-2002
DEFINITION Sequence 2 from Patent WO0192578.
ACCESSION AX323410
VERSION
KEYWORDS
SOURCE
ORGANISM human.
Human sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Roninson, I.B., Dokmanovic, M. and Chang, B.D.
TITLE Reagents and methods for identifying and modulating expression of genes regulated by retinoids

JOURNAL Patent: WO 0192578-A 2 06-DEC-2001;
Board of Trustees of the University of Illinois (US)
FEATURES Location/Qualifiers
source 1.1553
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Beta IG-H3 promoter NCBI acc. number AC004503.1"

BASE COUNT 305 a 490 c 384 g 374 t
ORIGIN

Query Match 100.0%; Score 1553; DB 6; Length 1553;
Best Local Similarity 100.0%; Pred No. 0;
Matches 1553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GACTCAGGAGTGTCCCAACACATCATCTACCTGGCAAGCTGCACTGCTGCTGCA 60
1 GACTCAGGAGTGTCCCAACACATCATCTACCTGGCAAGCTGCACTGCTGCTGCA 60
61 TTCTGAAACATGGACCATCTGCTGCAATGTCAGACACAAACACCTTCAATTCCT 120
61 TTCTGAAACATGGACCATCTGCTGCAATGTCAGACACAAACACCTTCAATTCCT 120
121 TGACCT 180
121 TGACCT 180
181 CTCCTACACCCCTGACATTTGCCAGGCTTCCCATCTGCTGCTGCTGCTGCTG 240
181 CTCCTACACCCCTGACATTTGCCAGGCTTCCCATCTGCTGCTGCTGCTGCTG 240
241 AGCTCCCAATGGCT 300
241 AGCTCCCAATGGCT 300
301 CACTGGCAGTGTCT 360
301 CACTGGCAGTGTCT 360
361 CACTGGCAGTGTCT 420
361 CACTGGCAGTGTCT 420
421 CGAGATCACTGATGTTGCGATGATGATGATGATGATGATGATGATGATGATG 480
421 CGAGATCACTGATGTTGCGATGATGATGATGATGATGATGATGATGATGATG 480
481 TTTCAGGCTCTAGAGGATGATGATGATGATGATGATGATGATGATGATGATG 540
481 TTTCAGGCTCTAGAGGATGATGATGATGATGATGATGATGATGATGATGATG 540
541 AGCCCCAGTCT 600
541 AGCCCCAGTCT 600
601 GCTTTTCAACAAGGCTTGAAGTCTCACTCATCTATTAATCACTGCTGCTCCAG 660
601 GCTTTTCAACAAGGCTTGAAGTCTCACTCATCTATTAATCACTGCTGCTCCAG 660
661 GCTAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 720
661 GCTAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 720
721 TACCATGTGACAAATTTGTTTACCTGCTTGAAGTCTCTCAAGGCAAGGCTTGG 780
721 TACCATGTGACAAATTTGTTTACCTGCTTGAAGTCTCTCAAGGCAAGGCTTGG 780
781 CAGATCTATCTGGGAGTCTGTTGAGAGATGTTTACGAGAGAGAGATGATAGT 840
781 CAGATCTATCTGGGAGTCTGTTGAGAGATGTTTACGAGAGAGAGATGATAGT 840
841 ACCAAGGTGAG 900
841 ACCAAGGTGAG 900

901 ACCTGCCCCACATTTTGGCGTTCTGTGATGATGATGATGATGATGATGATGAT 960
901 ACCTGCCCCACATTTTGGCGTTCTGTGATGATGATGATGATGATGATGATGAT 960
961 GAGATGACAGAGTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTT 1020
961 GAGATGACAGAGTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTT 1020
1021 CAGCTCTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTT 1080
1021 CAGCTCTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTT 1080
1081 CTCCCTACACCCCTCTCTGAGTGTGTTGAGCTCAAGATCAGTGAAGGATCTG 1140
1081 CTCCCTACACCCCTCTCTGAGTGTGTTGAGCTCAAGATCAGTGAAGGATCTG 1140
1141 CCCCCAGATGACAG 1200
1141 CCCCCAGATGACAG 1200
1201 TTGGAACAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
1201 TTGGAACAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
1261 GAGGAG 1320
1261 GAGGAG 1320
1321 CACAGTGGGAG 1380
1321 CACAGTGGGAG 1380
1381 AGGCGGCGCGCTCTCTCTGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
1381 AGGCGGCGCGCTCTCTCTGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
1441 TCCGCTGCGAGCTTACTTAACCTGAGGCGGCGGCGGCGGCGGCGGCGGCGG 1500
1441 TCCGCTGCGAGCTTACTTAACCTGAGGCGGCGGCGGCGGCGGCGGCGGCGG 1500
1501 CCGCGCGCTTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1553
1501 CCGCGCGCTTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1553

RESULT 2
AC004503
LOCUS AC004503 76007 bp DNA linear PRI 30-MAR-1998
DEFINITION Homo sapiens chromosome 5, p1 clone 1354A7 (LBNL H47), complete
sequence.
ACCESSION AC004503 AC002223 AC002225 AC001513 AC002226 AC001514
AC002222 AC001516 AC001515 AC001511 AC001510 AC001512 AC002221
AC002224
VERSION AC004503.1 GI:2996635
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 76007)
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,
Davis,C.A., Kadner,K., Miguel,T., Pitluck,S., Pollard,M.,
Rojeski,H., Subramanian,S. and Martin,C.H.
TITLE Large Scale Sequence Analysis and Annotation with the Sequence
Comparison Analysis (SCAN) System
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 76007)
Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,

TITLE Davis, C.A., Kadner, K., Miguel, T., Pitluck, S., Pollard, M.,
Rojski, H., Subramanian, S. and Martin, C.H.
JOURNAL Submitted (30-MAR-1998) Human Genome Center, DOE Joint Genome
Institute, Lawrence Berkeley National Laboratory, MS 74-157,
Berkeley, CA 94720, U.S.A.
COMMENT Sequence submitted by:
DOE Joint Genome Institute.
FEATURES Location/Qualifiers
1..76007
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/map="5q"
/clone="1354A7"
/note="LBNL H47"
complement(1..39820)
/note="99% identity AC002428"
repeat_region complement(92..1315)
/rpt_family="Alu"
repeat_region 1355..1444
/rpt_family="MSTC"
1558..1705
/rpt_family="MSTB"
complement(2377..5032)
/rpt_family="L1"
repeat_region complement(6287..6744)
/rpt_family="Alu"
complement(8102..9558)
/rpt_family="L1"
misc_feature complement(9587..9700)
/note="GRAIL 2 excellent exon, frame 0"
repeat_region 12179..12494
/rpt_family="MLT1a"
15197..15345
/rpt_family="MLT1a"
15425..15597
/rpt_family="MLT1a"
complement(17298..17716)
/rpt_family="L1"
repeat_region complement(18631..18813)
/rpt_family="MT1e"
complement(18989..19025)
/rpt_family="MLT1"
19607..19636
/note="(GT)15"
/rpt_type=tandem
/rpt_unit=GT
misc_feature complement(23378..23472)
/note="GRAIL 2 excellent exon, frame 0"
repeat_region complement(23777..24117)
/rpt_family="Alu"
24513..24586
/rpt_family="MER68B"
complement(2785..27918)
/rpt_family="MIR2"
misc_feature complement(28314..28444)
/note="GRAIL 2 excellent exon, frame 0"
repeat_region 29788..30276
/rpt_family="Alu"
30242..30272
/note="(A)31"
/rpt_type=tandem
/rpt_unit=A
repeat_region 30483..30800
/rpt_family="MER33"
complement(32562..33906)
/rpt_family="L1"
repeat_region complement(35700..35884)
/rpt_family="MER5"
complement(36451..36650)
/rpt_family="MIR"
repeat_region complement(37382..37458)

repeat_region /rpt_family="MIR2"
complement(41227..41588)
/rpt_family="THE1"
41633..41763
/rpt_family="MIR"
42050..42204
/note="GRAIL 2 excellent exon, frame 0"
44455..44591
/note="GRAIL 2 excellent exon, frame 2"
complement(46368..46472)
/rpt_family="MIR"
48908..>75715
/gene="BIGH3"
/note="transforming growth factor-beta induced gene
product (BIGH3)"
join(48955..49088,53662..53760,63957..64021,66234..66394,
66750..66914,67173..67319,69338..69479,72806..73018,
73842..73979,74615..74760,75579..>75715)
/gene="BIGH3"
/note="transforming growth factor-beta induced gene
product (BIGH3)"
/evidence=experimental
/product="BIGH3"
/protein_id="AAC08449.1"
/db_xref="GI:2996636"
/translation="MALFVRLALALALGPAATLAGPAKSPYOLVLRGRHOH
GNVAVOKVIGFNKRYFTNGKQWQKICGKSTVSYSCGGEVPEKCPALP
LNVLETGVSSTQLTQDTEKRPMEGRGSTTTPASNEANASIPAEYLSLV
SVNIEILNALRIHAYGRVLTDELHGHTLSMKNQNSIQIHHTPNGIVTNCALC
KADHAATNGVHHIDKVISFTNNIQIIEIDVPEFTLAAVAASLNTMLRNGCYT
LAPNEAPEKIPSETLNRILIDPEALRDLNNHILKSAACAALVAGLSVELEST
LEVGCSGDMLTNGKAIISNKDILATNGVIYHIDELIPDAKTELEAESPSTY
DLFRQAGLNLSGSEKRLFLAPINSEVFGDGPIDAHNRILLRHHITDOLASKY
HGCTLETTGSKRLRVYRNSLSCIENSCLAHDKRRTGLTFMDRVLPMPGTADV
LKGNRF"
55224..55552
/rpt_family="Alu"
58141..58334
/rpt_family="MLT1"
58303..58632
/rpt_family="Alu"
58629..58886
/rpt_family="MLT1"
60208..60533
/rpt_family="MER1B"
complement(62232..62307)
/rpt_family="MIR"
63480..63559
/rpt_family="MIR"
complement(63619..63756)
/note="GRAIL 2 excellent exon, frame 0"
/rpt_family="MER2"
65024..65474
/rpt_family="Alu"
65447..65474
/note="(AAAT)7"
/rpt_type=tandem
/rpt_unit=AAAT
repeat_region 69938..70262
/rpt_family="Alu"
71056..71368
/rpt_family="Alu"
71443..71590
/gene="BIGH3"
/note="transforming growth factor-beta induced gene
product (BIGH3)"
71652..71691
/gene="BIGH3"
75793..76002
/note="GRAIL 2 excellent exon, frame 0"
/note="100% identity B43847 BAC end sequence"
BASE COUNT 20153 a 16885 c 17588 g 21381 t

ORIGIN

Query Match	99.38;	Score 1542;	DB 9;	Length 76007;
Best Local Similarity	99.98;	Pred. No. 0;		
Matches 1553; Conservative	0;	Mismatches 0;	Indels 1;	Gaps 1

OY	1	GACGAGGGTGGCCAAACCACTCAATTCACGGGGAAGCTCAGCTCAGTGGCTCA	60
Db	47401	GACTCAGGGGTGCCAAACACTCATCTTACCTGGCAAGCTCAGCTCTCATATGGCTCTA	47460
OY	61	TTTCGAACATGGCACCACATCA-CTGCTGCATATGTCCAGAACACAAACCCCTACATATCC	119
Db	47461	TTTCGAACATGGCACCACATCACTCTGCTGCATATGTCCAGAACACAAACCCCTACATATCC	47520
OY	120	TTTGACTCTCTCTTCTCCCTCTCTCCCTGTATACAGACTCCAAATCTATTTGAGCATATTA	179
Db	47521	TTTGACTCTCTCTTCTCCCTCTCTCCCTGTATACAGACTCCAAATCTATTTGAGCATATTA	47580
OY	180	CCCTCTACACCCCTCACATTTTGGCCAGCCCTTCCCATCTGCTCTCTACACATAGTTC	239
Db	47581	CCCTCTACACCCCTCACATTTTGGCCAGCCCTTCCCATCTCTGCTCTCTACACATAGTTC	47640
OY	240	AAGCTCTCCCATGGTCCCTTCCTGGTACTGTTCTCTTGGCTCCTTAAGCCTCTCATG	299
Db	47641	AAGCTCTCCCATGGTCCCTTCCTGGTACTGTTCTCTTGGCTCCTTAAGCCTCTCATG	47700
OY	300	ACACTGGCCATGTACTTGGCTCCACCCATCACCCGCTAGGCTCTTAGTGGAGTCTGGG	359
Db	47701	ACACTGGCCATGTACTTGGCTCCACCCATCACCCGCTAGGCTCTTAGTGGAGTCTGGG	47760
OY	360	CCCTGCTACCTTCACCCCTCTTCTCTTACCTTACCTCACCTCCCTTGGTTCAGCA	419
Db	47761	CCCTGCTACCTTCACCCCTCTTCTCTTACCTTACCTCACCTCCCTTGGTTCAGCA	47820
OY	420	ACCAAGATTACTTGAATTCGTGATGATCATGCTCAGTTTACCTGATTAATCATTTTCAT	479
Db	47821	ACCAAGATTACTTGAATTCGTGATGATCATGCTCAGTTTACCTGATTAATCATTTTCAT	47880
OY	480	CTTTCAGGCCCTCAGAGGAGGTATCACCCGTACAGGCGCAGAGGCGCTCTTACAGTCCCA	539
Db	47881	CTTTCAGGCCCTCAGAGGAGGTATCACCCGTACAGGCGCAGAGGCGCTCTTACAGTCCCA	47940
OY	540	AAGCCCGCAGCTCTTCATGGAAATCATTTGGGCTACGGATCTTCCCAAAATTTGG	599
Db	47941	AAGCCCGCAGCTCTTCATGGAAATCATTTGGGCTACGGATCTTCCCAAAATTTGG	48000
OY	600	AGCTTTTTCACAAAAGGGGTTTAGTGTCTACTACTTCTATTAATCCATCTGTGTCTCCAG	659
Db	48001	AGCTTTTTCACAAAAGGGGTTTAGTGTCTACTACTTCTATTAATCCATCTGTGTCTCCAG	48060
OY	660	GGCTAGCAGTGCCTCAAGTAACTGACAGGTGATTAAATAGATGCTTGGGTAAAGTACACTCT	719
Db	48061	GGCTAGCAGTGCCTCAAGTAACTGACAGGTGATTAAATAGATGCTTGGGTAAAGTACACTCT	48120
OY	720	TTTACCATGTGCAATTTGGTTTACCTGCCCTTGAGCTCTCCAGGGCAGGACTCTTGCTTT	779
Db	48121	TTTACCATGTGCAATTTGGTTTACCTGCCCTTGAGCTCTCTCCAGGGCAGGACTCTTGCTTT	48180
OY	780	GCAAAATCTATCTGCGAGGTACTGTGTGAGAGATGTTTACTGTAAGAAGGAATGAATTAG	839
Db	48181	GCAAAATCTATCTGCGAGGTACTGTGTGAGAGATGTTTACTGTAAGAAGGAATGAATTAG	48240
OY	840	TACCAAGGTGAGGACCCCAACCTTCCCAAGGGCTCCAAAAGCACCTTAGAGCCCAACA	899
Db	48241	TACCAAGGTGAGGACCCCAACCTTCCCAAGGGCTCCAAAAGCACCTTAGAGCCCAACA	48300
OY	900	AACCTGCCCCACATTTTGGCGTTCTGTGTGATCACAGATTTACTCATCTGTCTTTCA	959
Db	48301	AACCTGCCCCACATTTTGGCGTTCTGTGTGATCACAGATTTACTCATCTGTCTTTCA	48360
OY	960	TGACAAATGACAGGTGGGTGGGGGGTGGAGAGATTTAGAATTTAGAGAGCTGGGGAAGGTGG	1019
Db	48361	TGACAAATGACAGGTGGGTGGGGGGTGGAGAGATTTAGAATTTAGAGAGCTGGGGAAGGTGG	48420

OY	1020	TCGCTCTGGGGGTCGAGAAACAAAGCTGATGAGGGCAGTGGTTCTGGAATCAGACAC	1079
OY	48421	TCAGCTCTGGGGGTCGAGAAACAAAGCTGATGAGGGCAGTGGTTCTGGAATCAGACAC	48480
OY	1080	CTTCCCTCAGCCCTCCCTGCAAGTGTGTTTGTAGCTCAAGATTCAGTAGAGGAATCTTCCG	1139
Db	48481	CTTCCCTCAGCCCTCCCTGCAAGTGTGTTTGTAGCTCAAGATTCAGTAGAGGAATCTTCCG	48540
OY	1140	GCCTCCAGCATGTCAGGACCGAAGCCCGAGACAGCTGTCCCTCAGTCCCAAGTCCCA	1199
Db	48541	GCCTCCAGCATGTCAGGACCGAAGCCCGAGACAGCTGTCCCTCAGTCCCAAGTCCCA	48600
OY	1200	TTTGGAGCAGCCACAGGAGGCGCTAAGGACCTATACCTCTGGTTTAGAGAAACATCTGG	1259
Db	48601	TTTGGAGCAGCCACAGGAGGCGCTAAGGACCTATACCTCTGGTTTAGAGAAACATCTGG	48660
OY	1260	CGAGGAGACAGGAGGAGGAGGCGTGCGACATGTAGGGCAGAGGCTGGGAAACTGAGGACGG	1319
Db	48661	CGAGGAGACAGGAGGAGGAGGCGTGCGACATGTAGGGCAGAGGCTGGGAAACTGAGGACGG	48720
OY	1320	GCAAGATGCGGAGAGCGGATGGGTGCCACAGGGCAGGCAAGGAGGCGCAGGATTTGGAGAGCC	1379
Db	48721	GCAAGATGCGGAGAGCGGATGGGTGCCACAGGGCAGGCAAGGAGGCGCAGGATTTGGAGAGCC	48780
OY	1380	CAGGAGGCGCGGCTCTCTTTCACAGGGACGGGCCACAGCTTCCCGGCGCTGGCTCCGCTCC	1439
Db	48781	CAGGAGGCGCGGCTCTCTTTCACAGGGACGGGCCACAGCTTCCCGGCGCTGGCTCCGCTCC	48840
OY	1440	CTTCCCTGCGACACTTACTTAACCTGCGCCCGGCGGCGGAGGCGCTCTCACTTCCCTGGA	1499
Db	48841	CTTCCCTGCGACACTTACTTAACCTGCGCCCGGCGGCGGAGGCGCTCTCACTTCCCTGGA	48900
OY	1500	GCAGCCCGATGGCCGTCGGTGCAGTACGCTGCGTCCGGGTGCTGCCTCC	1553
Db	48901	GCAGCCCGATGGCCGTCGGTGCAGTACGCTGCGTCCGGGTGCTGCCTCC	48954

[illegible]

*	132177	137419:	contig of 5243 bp in length
*	137420	137439:	gap of unknown length
*	137440	142254:	contig of 4815 bp in length
*	142255	142274:	gap of unknown length
*	142275	149021:	contig of 6747 bp in length
*	149022	149041:	gap of unknown length
*	149042	152821:	contig of 3780 bp in length
*	149044	152842:	gap of unknown length
*	152842	157555:	contig of 4714 bp in length
*	157556	157375:	gap of unknown length
*	157576	160334:	contig of 2739 bp in length
*	160315	160334:	gap of unknown length
*	160335	164544:	contig of 4210 bp in length
*	164545	164564:	gap of unknown length
*	164565	166647:	contig of 2083 bp in length
*	166648	166657:	gap of unknown length
*	166658	166833:	contig of 3166 bp in length
*	169834	169853:	gap of unknown length
*	169834	173590:	contig of 3737 bp in length
*	173591	173610:	gap of unknown length
*	173611	173614:	contig of 4554 bp in length
*	178165	178184:	gap of unknown length
*	178185	180900:	contig of 2716 bp in length
*	180901	180920:	gap of unknown length
*	180921	182482:	contig of 1562 bp in length
*	182483	182502:	gap of unknown length
*	182503	185289:	contig of 2787 bp in length
*	185290	185309:	gap of unknown length
*	185310	186932:	contig of 1623 bp in length
*	186933	186952:	gap of unknown length
*	186953	188165:	contig of 1213 bp in length
*	188166	188185:	gap of unknown length
*	188186	189092:	contig of 907 bp in length
*	189093	189112:	gap of unknown length
*	189113	190678:	contig of 1566 bp in length
*	190679	190698:	gap of unknown length
*	190699	192877:	contig of 2179 bp in length
*	192878	192897:	gap of unknown length
*	192898	194214:	contig of 1317 bp in length
*	194215	194234:	gap of unknown length
*	194235	195535:	contig of 1301 bp in length
*	195536	195555:	gap of unknown length
*	195556	196407:	contig of 852 bp in length
*	196408	196427:	gap of unknown length
*	196428	197671:	contig of 1244 bp in length
*	197672	197691:	gap of unknown length
*	197692	198960:	contig of 1269 bp in length
*	198961	198980:	gap of unknown length
*	198981	200542:	contig of 1562 bp in length
*	200543	200562:	gap of unknown length
*	200563	201492:	contig of 930 bp in length
*	201493	202368:	contig of 856 bp in length
*	202368	202388:	gap of unknown length
*	202389	203978:	contig of 1590 bp in length
*	203979	203998:	gap of unknown length
*	203999	205308:	contig of 1310 bp in length
*	205309	205328:	gap of unknown length
*	205329	207420:	contig of 2092 bp in length

```
misc_feature 1. 63778
              /note="assembly_name:Contig130"
misc_feature 63799. 96902
              /note="assembly_name:Contig129
misc_feature 96923. 107488
              /note="assembly_name:Contig128"
```

```

misc_feature      107509. .119219 /note="assembly_name:Contig127"
misc_feature      119240. .132156 /note="assembly_name:Contig126"
clone_end="77"
vector_side:right"
132177. .137419
/note="assembly_name:Contig125"
137440. .142254
/note="assembly_name:Contig124"
142275. .145021
/note="assembly_name:Contig123"
149042. .152821
/note="assembly_name:Contig122"
152842. .157555
/note="assembly_name:Contig121"
157576. .160314
/note="assembly_name:Contig120"
160335. .164544
/note="assembly_name:Contig119"
164565. .166647
/note="assembly_name:Contig118"
166668. .169833
/note="assembly_name:Contig117"
clone_end:Sp6
vector_side:right"
169854. .173590
/note="assembly_name:Contig116"
173611. .178164
/note="assembly_name:Contig115"
178185. .180900
/note="assembly_name:Contig114"
180921. .182482
/note="assembly_name:Contig113"
182503. .185289
/note="assembly_name:Contig112"
185310. .186932
/note="assembly_name:Contig111"
186953. .188165
/note="assembly_name:Contig110"
188186. .189092
/note="assembly_name:Contig109"
189113. .190678
/note="assembly_name:Contig108"
190699. .192877
/note="assembly_name:Contig107"
192898. .194214
/note="assembly_name:Contig106"
194235. .195535
/note="assembly_name:Contig105"
195536. .196407
/note="assembly_name:Contig104"
196428. .197671
/note="assembly_name:Contig103"
197692. .198960
/note="assembly_name:Contig102"
198981. .200542
/note="assembly_name:Contig101"
200563. .201492
/note="assembly_name:Contig100"
201513. .202368
/note="assembly_name:Contig99"
202389. .203978
/note="assembly_name:Contig98"
203999. .205308
/note="assembly_name:Contig97"
205329. .207420

Query Match      4.9% ; Score 75.8 ; DB 2 ; Length 207420 ;
Best Local Similarity 52.9% ; Pred.No. 6.4e-08 ;
Matches 164 ; Conservative 0 ; Mismatches 144 ; Indels 2 ; Gaps 1 ;

```


BASE COUNT	39694	a	32483	c	32177	g	40192	t	6737	others
ORIGIN										
Query Match	4.7%; Score 73; DB 2; Length 151283;									
Best Local Similarity	51.9%; Pred. No. 3.1e-07;									
Matches 154; Conservative	0; Mismatches 143; Indels 0; Gaps 0;									
OY 1248	GGAAGACTGTGGCGGAGAGAGAGGAGGAGGAGCGTGCACGTGAGAGCCAGAGCTGGGAA	1307								
Db 26762	GGGGGNNGGGGGGGGGGGNGGGGGGCGNCGGGGGGGCGGNNGGGGGGGGGGNGGGGGC	26821								
OY 1308	AACGACACAGGGCACAGTCGGGAGCGGGTGTTGCCAAGGCAGCCAGGGCGGCAGCG	1367								
Db 26822	GGGNGGGGGGGGGGGC	26881								
OY 1368	GTTGGAGAGGGGACGAGCGGGCCCCCTCTTTCACAGGGCGGGCCAGATTCCCAGCCCC	1427								
Db 26882	CGGG	26941								
OY 1428	GGCGTCGCCGTCCCTCCCGCTCGCAGCTTACCTTAACCTGCGCCGGGCGGCGAGGCGCTCT	1487								
Db 26942	GGCCCCCGCGG	27001								
OY 1488	CACCTCCCTTGAGCGCCCGCTTGCCCGTGGCTGCTAGCTGCTCGGTGCGCGCTGG	1544								
Db 27002	CCGGCCCCCGG	27058								
RESULT 7										
AK094055	2583 bp mRNA linear PRI 15-JUL-2002									
LOCUS	Homo sapiens CDNA FLJ36736 fls, clone UTERU2013048, highly similar	TO TRANSFORMING GROWTH FACTOR-BETA INDUCED PROTEIN IG-H3 PRECURSOR								
ACCESSION	AK094055									
VERSION	AK094055.1 GI:21753035									
KEYWORDS	oligo capping; fts (full insert sequence).									
SOURCE	Homo sapiens uterus cdna to mRNA, clone_lib=UTERU2									
ORGANISM	clone:UTERU2013048.									
REFERENCE	Homo sapiens									
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
1	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.									
Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R.,										
Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,										
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,										
Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,										
Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,										
Oshima,A., Suzuki,Y., Sugano,S., Nagahari,K., Masuno,Y., Nagai,K.										
and Isogai,T.										
NEDO human cdna sequencing project										
unpublished										
2 (bases 1 to 2583)										
Isogai,T. and Yamamoto,J.										
Direct Submission										
Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7										
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan										
(E-mail:genomics@hri.co.jp; tel:81-438-52-3975, Fax:81-438-52-3986										
NEDO human cdna sequencing project supported by Ministry of										
Economy, Trade and Industry of Japan; cdna full insert sequencing:										
Research Association for Biotechnology (RAB); cdna library										
construction: Helix Research Institute (HRI) (supported by Japan										
Key Technology Center etc.); 5'- & 3'-end one pass sequencing; RAB,										
HRI, and Biotechnology Center, National Institute of Technology and										
Evaluation; clone selection for full insert sequencing; HRI and										
RAB; annotation: HRI and RAB.										
Location/Qualifiers										
1.. 2583										
/organism="Homo sapiens"										
/db_xref="taxon:9606"										
/clone=UTERU2013048"										
/tissue="uterus"										

TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL

Newton, N., Nguyen, N., Norris, S., Nwackelam, O., Okunolu, G.,
Olanupinsagbon, A., Pal, S., Parks, K., Pasternak, S., Paul, H.,
Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A.,
Popovic, D., Primus, E., Pu, L., Puazo, M., Quiroz, J., Rachlin, E.,
Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y.,
Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A.,
Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Saverly, G., Scherer, S.,
Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A.,
Sisson, I., Sitter, C. D., Smajs, D., Sneed, A., Sodergren, E.,
Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A.,
Starek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S.,
Tiney, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D.,
Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J.,
Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R.,
Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S.,
Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X.,
Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R.,
Holt, R. A., Smith, H. O., Weinstein, G. and Gibbs, R. A.

Direct Submission
Unpublished
2 (bases 1 to 100740)

Rat Genome Sequencing Consortium.

Submitted (18-AUG-2002)

Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project information

Center project name: KAOA

Center clone name: CH230-62N13

Summary Statistics

Sequencing vector: Plasmid

Chemistry: Dye-terminator Big Dye: 100% of reads

Assembly program: Phrap, version 0.990329

Consensus quality: 47096 bases at least Q40

Consensus quality: 50680 bases at least Q30

Consensus quality: 54217 bases at least Q20

NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

NOTE: This is a 'working draft' sequence. It currently

consists of 52 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 1330: contig of 1330 bp in length
* 1331 1430: gap of unknown length
* 1431 2453: contig of 1023 bp in length
* 2454 2554: gap of unknown length
* 2554 3619: contig of 1066 bp in length
* 3620 3720: gap of unknown length
* 3720 5059: contig of 1340 bp in length
* 5060 5159: gap of unknown length
* 5160 6654: contig of 1495 bp in length
* 6655 7927: gap of unknown length
* 7927 8026: gap of unknown length
* 8027 9404: contig of 1378 bp in length
* 9405 9504: gap of unknown length
* 9505 10747: contig of 1243 bp in length
* 10748 10847: gap of unknown length
* 10848 12238: contig of 1391 bp in length
* 12239 12338: gap of unknown length
* 12339 13760: contig of 1422 bp in length
* 13761 13860: gap of unknown length
* 13861 15505: contig of 1645 bp in length

15506 15605: gap of unknown length
* 15606 17187: contig of 1582 bp in length
* 17188 17287: gap of unknown length
* 17288 18925: contig of 1638 bp in length
* 18926 19025: gap of unknown length
* 19026 20382: contig of 1337 bp in length
* 20383 20483: gap of unknown length
* 20483 21644: contig of 1162 bp in length
* 21645 21745: gap of unknown length
* 21745 23161: contig of 1417 bp in length
* 23162 23261: gap of unknown length
* 23262 24853: contig of 1552 bp in length
* 24854 24954: gap of unknown length
* 24954 25987: contig of 1034 bp in length
* 25988 26087: gap of unknown length
* 26088 27465: contig of 1378 bp in length
* 27466 27565: gap of unknown length
* 27566 29172: contig of 1607 bp in length
* 29173 29272: gap of unknown length
* 29273 30410: contig of 1138 bp in length
* 30411 30510: gap of unknown length
* 30511 32419: contig of 1909 bp in length
* 32420 32519: gap of unknown length
* 32520 34534: contig of 2015 bp in length
* 34535 34634: gap of unknown length
* 34635 36068: contig of 1434 bp in length
* 36069 36168: gap of unknown length
* 36169 37561: contig of 1393 bp in length
* 37562 37661: gap of unknown length
* 37662 39373: contig of 1712 bp in length
* 39374 39473: gap of unknown length
* 39474 41883: contig of 2410 bp in length
* 41884 41983: gap of unknown length
* 41984 43914: contig of 1931 bp in length
* 43915 44014: gap of unknown length
* 44015 45567: contig of 1553 bp in length
* 45568 45667: gap of unknown length
* 45668 47087: contig of 1420 bp in length
* 47088 47187: gap of unknown length
* 47188 48713: contig of 1526 bp in length
* 48714 48813: gap of unknown length
* 48814 50876: contig of 2063 bp in length
* 50877 50976: gap of unknown length
* 50977 52113: contig of 1137 bp in length
* 52114 52213: gap of unknown length
* 52214 53320: contig of 1107 bp in length
* 53321 53420: gap of unknown length
* 53421 55496: contig of 2076 bp in length
* 55497 55596: gap of unknown length
* 55597 56811: contig of 1215 bp in length
* 56812 56911: gap of unknown length
* 56912 58502: contig of 1591 bp in length
* 58503 58602: gap of unknown length
* 58603 59807: contig of 1205 bp in length
* 59808 59907: gap of unknown length
* 59908 62370: contig of 2363 bp in length
* 62371 62371: gap of unknown length
* 62371 64542: contig of 2172 bp in length
* 64543 64642: gap of unknown length
* 64643 66802: contig of 2160 bp in length
* 66803 66902: gap of unknown length
* 66903 70141: contig of 3239 bp in length
* 70142 70241: gap of unknown length
* 70242 72440: contig of 2199 bp in length
* 72441 72540: gap of unknown length
* 72541 74292: contig of 1732 bp in length
* 74293 74392: gap of unknown length
* 74393 76503: contig of 2111 bp in length
* 76504 76603: gap of unknown length
* 76604 79013: contig of 2410 bp in length
* 79014 79113: gap of unknown length
* 79114 81669: contig of 2556 bp in length
* 81670 81769: gap of unknown length

Massey E., Mawhney E., McLeod M.P., Meador M., Mel G., Metzker M.,
Miner G., Miner Z., Mitchell T., Mohabbat K., Morgan M., Morris S.,
Moser M., Neal D., Newton J., Newton N., Nguyen A., Nguyen N.,
Nguyen N., Nickerson E., Nwokoko S., Oguh M., Okunolu G.,
Oragunye N., Oviedo R., Pace A., Payton B., Peery J., Perez L.,
Peters L., Pickens R., Primus E., Pu L.L., Quiles M., Ren Y.,
Rives M., Rojts A., Rojudoan L., Rolfe M., Ruiz S., Saverly G.,
Scherer S., Scott G., Shen H., Shooshari N., Sisson T.,
Sodergren E., Sonaike T., Sparks A., Stanley H., Stone H.,
Sutton A., Sytek A., Tabor P., Tamerisa A., Tamerisa K., Tang H.,
Taneey J., Taylor C., Taylor T., Telford B., Thomas N., Thomas S.,
Usmani K., Vasquez L., Vera V., Villalón D., Vinton R., Wang Q.,
Wang S., Ward-Moore S., Warren R., Washington C., Wallington S.,
Williams G., Williamson A., Wleczek R., Wooden S., Worley K.,
Wu C., Wu Y., Wu Y.F., Zhou J., Zorrilla S., Nelson D.,
Weinstock G. and Gibbs R.

Direct Submission
Unpublished
2 (bases 1 to 119989)
Worley K.C.

Direct Submission
Submitted (27-APR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 119989)
Worley K.C.

Direct Submission
Submitted (18-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 14, 2002 this sequence version replaced gi:20429765.

----- Genome Center -----
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information -----
Center project name: GVEP
Center clone name: CH230-415B14

----- Summary Statistics -----
Sequencing vector: Plasmid:
Chemistry: Dye-terminator Big Dye 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 71249 bases at least Q40
Consensus quality: 76269 bases at least Q30
Consensus quality: 79649 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 48 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1014: contig of 1014 bp in length
* 1015 1114: gap of unknown length
* 1115 2599: contig of 1485 bp in length
* 2600 2699: gap of unknown length
* 2700 4257: contig of 1558 bp in length
* 4258 4357: gap of unknown length
* 4358 5720: contig of 1363 bp in length
* 5721 5820: gap of unknown length
* 5821 7353: contig of 1533 bp in length
* 7354 7453: gap of unknown length
* 7454 8707: contig of 1254 bp in length
* 8708 8807: gap of unknown length
* 8808 9962: contig of 1155 bp in length
* 9963 10062: gap of unknown length
* 10063 11789: contig of 1727 bp in length
* 11790 11889: gap of unknown length

11890 12898: contig of 1009 bp in length
* 12899 12998: gap of unknown length
* 12999 14485: contig of 1487 bp in length
* 14486 14585: gap of unknown length
* 14586 15746: contig of 1161 bp in length
* 15747 15846: gap of unknown length
* 15847 15902: contig of 1056 bp in length
* 15903 17002: gap of unknown length
* 17003 18295: contig of 1292 bp in length
* 18296 18395: gap of unknown length
* 18396 19731: contig of 1337 bp in length
* 19732 19832: gap of unknown length
* 19833 20840: contig of 1009 bp in length
* 20841 20941: gap of unknown length
* 20942 22881: contig of 1940 bp in length
* 22882 22981: gap of unknown length
* 22982 24470: contig of 1490 bp in length
* 24471 24570: gap of unknown length
* 24571 26235: contig of 1665 bp in length
* 26236 26335: gap of unknown length
* 26336 27539: contig of 1204 bp in length
* 27540 27640: gap of unknown length
* 27641 29666: contig of 2027 bp in length
* 29667 29766: gap of unknown length
* 29767 32144: contig of 2378 bp in length
* 32145 32244: gap of unknown length
* 32245 34551: contig of 2307 bp in length
* 34552 34651: gap of unknown length
* 34652 36879: contig of 2228 bp in length
* 36880 36979: gap of unknown length
* 36980 38695: contig of 1716 bp in length
* 38696 38795: gap of unknown length
* 38796 40177: contig of 1381 bp in length
* 40178 40276: gap of unknown length
* 40277 42504: contig of 2228 bp in length
* 42505 42604: gap of unknown length
* 42605 44375: contig of 1771 bp in length
* 44376 44475: gap of unknown length
* 44476 46073: contig of 1598 bp in length
* 46074 46174: gap of unknown length
* 46175 47658: contig of 1485 bp in length
* 47659 50216: gap of unknown length
* 50217 50316: contig of 2458 bp in length
* 50317 52271: contig of 1955 bp in length
* 52272 52371: gap of unknown length
* 52372 54575: contig of 2204 bp in length
* 54576 54675: gap of unknown length
* 54676 57590: contig of 2915 bp in length
* 57591 59891: gap of unknown length
* 59892 59991: contig of 2201 bp in length
* 59992 61347: gap of unknown length
* 61348 61447: contig of 1356 bp in length
* 61448 64459: gap of unknown length
* 64460 64559: contig of 3012 bp in length
* 64560 66054: gap of unknown length
* 66055 68054: contig of 3395 bp in length
* 68055 71223: gap of unknown length
* 71224 71323: contig of 3169 bp in length
* 71324 73860: gap of unknown length
* 73861 73961: contig of 2537 bp in length
* 73962 77898: gap of unknown length
* 77899 80364: contig of 3938 bp in length
* 80365 80464: gap of unknown length
* 80465 85763: gap of unknown length
* 85764 85863: contig of 5299 bp in length
* 85864 90800: gap of unknown length
* 90801 90900: contig of 4937 bp in length
* 90901 96288: gap of unknown length
* 96289 96388: contig of 5388 bp in length
* 96389 100772: gap of unknown length
* 100773 100772: contig of 4384 bp in length

Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Franz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, R., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hognes, M., Holway, C., Hollins, B., Homst, F., Howard, S., Huber, D., Huix, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolyet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lien, C., Liu, J., Liu, W., Loussaged, H., Lozano, R., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maneswari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mel, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokwenkwo, S., Oguh, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojudo, I., Rolfe, M., Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shoshitari, N., Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S., Williams, G., Williamson, A., Wleczek, R., Woodson, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 147622)
Worley, K.C.

Direct Submission
Submitted (13-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 147622)
Worley, K.C.

Direct Submission
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jul 12, 2002 this sequence version replaced gl:18846979.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: GOGF
Center clone name: CH230-240E10

----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap version 0.990329
Consensus quality: 99292 bases at least Q40
Consensus quality: 103354 bases at least Q30
Consensus quality: 106749 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 59 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1040: contig of 1040 bp in length
* 1041 1140: gap of unknown length
* 1141 2270: contig of 1130 bp in length

2271 2370: gap of unknown length
* 2371 3422: contig of 1052 bp in length
* 3423 3522: gap of unknown length
* 3523 4533: contig of 1011 bp in length
* 4534 4633: gap of unknown length
* 4634 6294: contig of 1661 bp in length
* 6295 6394: gap of unknown length
* 6395 7900: contig of 1506 bp in length
* 7901 8000: gap of unknown length
* 8001 9008: contig of 1008 bp in length
* 9009 9108: gap of unknown length
* 9109 10149: contig of 1041 bp in length
* 10150 10249: gap of unknown length
* 10250 11661: contig of 1412 bp in length
* 11662 11761: gap of unknown length
* 11762 13213: contig of 1452 bp in length
* 13214 13313: gap of unknown length
* 13314 14852: contig of 1539 bp in length
* 14853 14952: gap of unknown length
* 14953 16042: contig of 1090 bp in length
* 16043 16142: gap of unknown length
* 16143 17532: contig of 1390 bp in length
* 17533 17633: gap of unknown length
* 17634 18933: contig of 1201 bp in length
* 18934 18933: gap of unknown length
* 18934 20069: contig of 1136 bp in length
* 20070 20169: gap of unknown length
* 20170 21285: contig of 1116 bp in length
* 21286 21385: gap of unknown length
* 21386 22422: contig of 1037 bp in length
* 22423 22522: gap of unknown length
* 22523 24767: contig of 2245 bp in length
* 24768 24867: gap of unknown length
* 24868 26530: contig of 1663 bp in length
* 26531 26630: gap of unknown length
* 26631 28662: contig of 2032 bp in length
* 28663 28762: gap of unknown length
* 28763 29989: contig of 1227 bp in length
* 29990 30089: gap of unknown length
* 30090 31439: contig of 1350 bp in length
* 31440 31539: gap of unknown length
* 31540 33774: contig of 2135 bp in length
* 33775 33774: gap of unknown length
* 33775 35587: contig of 1813 bp in length
* 35588 35687: gap of unknown length
* 35689 38101: contig of 2414 bp in length
* 38102 38201: gap of unknown length
* 38202 39526: contig of 1325 bp in length
* 39527 39626: gap of unknown length
* 39627 41384: contig of 1758 bp in length
* 41385 41484: gap of unknown length
* 41485 43477: contig of 1993 bp in length
* 43478 43577: gap of unknown length
* 43578 44733: contig of 1156 bp in length
* 44734 44833: gap of unknown length
* 44834 46427: contig of 1594 bp in length
* 46428 46527: gap of unknown length
* 46529 48229: contig of 1702 bp in length
* 48230 48329: gap of unknown length
* 48330 49559: contig of 1230 bp in length
* 49560 49659: gap of unknown length
* 49660 50956: contig of 1297 bp in length
* 50957 51056: gap of unknown length
* 51057 53501: contig of 2445 bp in length
* 53502 53601: gap of unknown length
* 53602 56724: contig of 3123 bp in length
* 56725 56824: gap of unknown length
* 56825 59743: contig of 2819 bp in length
* 59744 59743: gap of unknown length
* 59744 61622: contig of 1879 bp in length
* 61623 61722: gap of unknown length
* 61723 64577: contig of 2855 bp in length
* 64578 64677: gap of unknown length

*	66678:	contig of 2567	bp in length
*	667245	gap of unknown	length
*	67345	contig of 1430	bp in length
*	68775	gap of unknown	length
*	68875	contig of 2059	bp in length
*	70934	gap of unknown	length
*	71034	contig of 1741	bp in length
*	72775	gap of unknown	length
*	72875	contig of 2886	bp in length
*	75761	gap of unknown	length
*	75861	contig of 2841	bp in length
*	76702	gap of unknown	length
*	78802	contig of 2668	bp in length
*	81490	gap of unknown	length
*	85093	contig of 3504	bp in length
*	85094	gap of unknown	length
*	88194	contig of 2693	bp in length
*	87863	gap of unknown	length
*	87963	contig of 3288	bp in length
*	91251	gap of unknown	length
*	91351	contig of 2955	bp in length
*	92946	gap of unknown	length
*	94046	contig of 3324	bp in length
*	97370	gap of unknown	length
*	97470	contig of 4035	bp in length
*	101505	gap of unknown	length
*	105466	contig of 3861	bp in length
*	105665	gap of unknown	length
*	105666	contig of 6143	bp in length
*	111709	gap of unknown	length
*	111808	contig of 4153	bp in length
*	111809	gap of unknown	length
*	11561:	contig of 4153	bp in length

Query Match	4.08;	Score 62.4;	DB 2;	Length 147622;
Best Local Similarity	45.3%;	Pred. No. 0.00013;		
Matches 156; Conservative	0;	Mismatches 188;	Indels 0;	Gaps 0

OY	1203	GGAAGCAGCCACGAGAGGGCTTAATACCCTTGTTGAGGAAGACTGTGGCGA	1262
Db	3892	GGGGGGGGGNNNGGGGGGGCGCGGGGNGNNGNNNNNNNGANGNGNGGGGGGGGGG	3951
OY	1263	GGGAGAGAGGAGGAGGAGGCTGTGCAGTAGAGGCCAAGGCTGGGAAATACTGACAGCGGCA	13222
Db	3952	GGGGGGGGGGGGGGGGGGGGGGMGGGGGGGGGGGGGGGGGGGGGGGGGGG	40111
OY	1323	CAGTGCAGGAGCGGGGTGTGCCAGGCGCAGCGAGGGCGCACGGGTTTGGAGAGCGCAG	13828
Db	4012	GGMGGNNNGG	40711
OY	1383	GGGGCCCCCTCTTTCACAGGCGCGGCCAACATTCCCAGCCCTTGCGCTCCCTCTC	14422
Db	4072	GCCCGMBCGCGGGGGCGCGGGCGGCCCGCCCGCCCGCCGACGAGCGGGCGCCG	41311
OY	1443	CCGCGTCAGACTTACTTAACCTTGCGCGGCGGCGGAGGCGCTTCACATTCCCTGAGACC	15022
Db	4132	GGCGGGGCGCGNCGCCGCCCGCGGGCCCGGGCGCCGGCGCCGCCGCCGCCGCCGCG	41911
OY	1503	GCCCCGTTGCCGCTCGGTGCGTAGCTCGCTCGGTGCGGCTGTC	1546
Db	4192	CGGCGCCGMCCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC	4235

	RESULT	LOCUS	DEFINITION
ACI19304	201650 bp	DNA	linear HTG 18-JUL-2007 Rattus norvegicus clone CH230-105FII, *** SEQUENCING IN PROGRESS ***, 71 unordered pieces.

ACCESSION	AC119304
VERSION	AC119304.3
KEYWORDS	GI:21746367
SOURCE	HTG: HTGS_PHASE1.
ORGANISM	Norway rat.
	<i>Rattus norvegicus</i>
	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae;

REFERENCE
AUTHORS

Rattus.
1 (bases 1 to 201650)

Muzny, D. M., Adams, C., Adio-Ofodola, B., Ali-osman, F. R., Allen, C., Alshrooks, S. L., Amarasinghe, H. C., Are, J. R., Ayele, M., Banks, T., Barberia, J., Benton, J., Blmage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Briteva, M., Brown, E., Brown, M., Bryant, N. P., Buhay, C., Burch, P., Burkett, C., Butrell, K. L., Byrd, N. C., Caron, T. F., Carter, M., Cavazos, S. R., Chacko, J., Chavez, C., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C. D., Cox, C., Coyle, M. D., Dathorne, S. R., David, R., Davila, M. L., Davis, C., Davy-Carroll, L., Dedetich, D. A., Delaney, K. R., Delgado, O., Dem, A. L., Ding, Y., Dinh, H. H., Douthwaite, R. J., Draper, H., Dugan-Rocha, S., Durbin, K. J., Earnhardt, C., Edgar, D., Edwards, C. C., Elhaj, C., Escotto, M., Falls, T., Farrington, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorell, J. H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, R., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Huliy, S., Hunne, J., Jackson, L. E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Kovah, J., Kovar, C., Kratoch, J., Kureshi, A., Landry, N., Leal, B., Lewis, L. C., Lewis, L., Li, J., Li, Z., Licharge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H., Lozado, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M. P., Medora, M., Mel, G., Metzger, M., Miner, G., Mier, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okunoye, G., Oragunye, N., Orvedo, R., Pock, A., Payton, B., Peety, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L. L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojoubokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shooshitari, N., Sisson, I., Sodegrien, E., Sonalike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Swatek, A., Tabors, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansley, J., Taylor, C., Taylor, T., Telford, A., Thomas, N., Thomas, S., Usmani, R., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S., Williams, G., Williamson, A., Wlasczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorilla, S., Nelson, D., Weinstock, G., and Gibbs, R.

JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 201650)
AUTHORS	Worley, K.C.
TITLE	Direct Submision
JOURNAL	Submitted (26-Apr-2002) Human Genome Sequencing Center, Departmentt of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA 3 (bases 1 to 201650)
REFERENCE	Worley, K.C.
AUTHORS	Direct Submision
TITLE	Submitted (18-Jul-2002) Human Genome Sequencing Center, Departmentt of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
JOURNAL	On Jul 14, 2002 this sequence version replaced g1:20429676.
COMMENT	

Center Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

Project Information
Center project name: GUHE
Center clone name: CH230-105F1

Summary Statistics
Sequencing vector: Plasmid:
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap, version 0.990329
Consensus quality: 134842 bases at least Q40
Consensus quality: 140086 bases at least Q30
Consensus quality: 146272 bases at least Q20

NOTE: Estimated insert size may differ from sequence length (see http://www.hscg.bcm.tmc.edu/docs/genbank_drift_data.html)

NOTE: This is a 'working draft' sequence. It currently consists of 71 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1	1003: contig of 1003 bp in length
1004	1103: gap of unknown length
1104	2327: contig of 1224 bp in length
2328	2427: gap of unknown length
2428	3518: contig of 1091 bp in length
3519	3618: gap of unknown length
3619	4827: contig of 1209 bp in length
4828	4927: gap of unknown length
4928	6262: contig of 1335 bp in length
6263	6362: gap of unknown length
6363	7820: contig of 1458 bp in length
7821	7920: gap of unknown length
7921	9354: contig of 1434 bp in length
9355	9454: gap of unknown length
9455	11161: contig of 1707 bp in length
11162	11261: gap of unknown length
11262	12792: contig of 1531 bp in length
12793	12892: gap of unknown length
12893	14188: contig of 1296 bp in length
14189	14288: gap of unknown length
14289	15494: contig of 1206 bp in length
15495	15594: gap of unknown length
15595	17269: contig of 1675 bp in length
17270	17369: gap of unknown length
17370	18796: contig of 1427 bp in length
18797	18896: gap of unknown length
18897	20055: contig of 1159 bp in length
20056	20155: gap of unknown length
20156	21448: contig of 1293 bp in length
21449	21548: gap of unknown length
21549	22801: contig of 1253 bp in length
22802	22901: gap of unknown length
22902	24220: contig of 1319 bp in length
24221	24320: gap of unknown length
24321	25523: contig of 1203 bp in length
25524	25623: gap of unknown length
25624	27392: contig of 1769 bp in length
27393	27492: gap of unknown length
27493	28513: contig of 1021 bp in length
28514	28613: gap of unknown length
28614	30125: contig of 1512 bp in length
30126	30225: gap of unknown length
30226	31426: contig of 1201 bp in length
31427	31526: gap of unknown length
31527	32792: contig of 1271 bp in length
32798	32897: gap of unknown length
32898	34018: contig of 1121 bp in length
34019	34118: gap of unknown length
34119	35597: contig of 1479 bp in length
35598	35697: gap of unknown length
35698	37881: contig of 2184 bp in length
37882	37981: gap of unknown length
37982	39480: contig of 1499 bp in length
39481	39580: gap of unknown length
39581	40952: contig of 1372 bp in length
40953	41052: gap of unknown length
41053	42722: contig of 1670 bp in length
42723	42822: gap of unknown length
42823	44396: contig of 1114 bp in length
44397	44036: gap of unknown length
44037	47045: contig of 3009 bp in length
47046	47145: gap of unknown length
47146	49294: contig of 2149 bp in length

[illegible]

RESULT 15
AC131133
LOCUS
DEFINITION
AC131133 203343 bp DNA linear HTG 17-AUG-2002
Rattus norvegicus clone CH230-202D8, *** SEQUENCING IN PROGRESS
***, 67 unordered pieces.
AC131133
VERSION
AC131133.1 GI:22296618
HTG: HTG_PHASE1.
SOURCE
Norway rat.
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE

AUTHORS

1 (bases 1 to 203343)

Muzny,D.,Marle., Metzker,M.,Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alspbrooks,S., Amla,A., Anguiano,D.,
Anyalebech,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Catroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinu,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Frazer,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregiorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulik,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowls,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenshew,L., Loulseged,H., Lozano,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapa,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M., McNeill,T., Meenen,E., Milosavljevic,A.,
Miner,G., Ming,E., Montemayor,J., Moore,S., Morgan,M., Morris,K.,
Morris,S., Muidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D.,
Newton,N., Nguyen,N., Norris,S., Nwokeneme,O., Okwunou,G.,
Olanunnsagoo,A., Pal,S., Parks,K., Pasternak,S., Paul,H.,
Perez,A., Perez,L., Pfankoch,C., Plopper,F., Poindexter,A.,
Popovic,D., Primus,E., Fu,L.-L., Puzo,M., Quito,J., Rachlin,E.,
Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y.,
Reuter,M., Richards,S., Riggs,F., Rivers,C., Rodkey,T., Rojas,A.,
Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,S., Scherer,S.,
Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A.,
Sisson,I., Sitter,C.D., Smaj,D., Sneed,A., Sodergren,E.,
Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A.,
Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S.,
Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villaseana,D.,
Walton,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,D.,
Warren,R., Wei,X., White,F., Williams,G., Willison,R., Wleczky,R.,
Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S.,
Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X.,
Zhao,S., Dunn,D., von Niederhausen,A., Weiss,R., Smith,D.R.,
Holt,R.A., Smith,H.O., Weinstein,G. and Gibbs,R.A.

JOURNAL

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

Center: Baylor College of Medicine

Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: GKHP
Center clone name: CH230-202D8
Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 142566 bases at least Q40
Consensus quality: 152085 bases at least Q30
Consensus quality: 159708 bases at least Q20
NOTE: Estimated insert size may differ from sequence length
(see <http://www.hgsc.bcm.tmc.edu/docs/genbank/draft.html>).
NOTE: This is a 'working draft' sequence. It currently
consists of 67 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1
1033: contig of 1033 bp in length
1034: gap of unknown length
1133: contig of 1188 bp in length
1134: gap of unknown length
2321: gap of unknown length
2421: gap of unknown length
2422: contig of 1181 bp in length
3602: gap of unknown length
3702: gap of unknown length
4833: contig of 1131 bp in length
4834: gap of unknown length
4933: gap of unknown length
6334: contig of 1401 bp in length
6335: gap of unknown length
6435: contig of 1776 bp in length
8210: gap of unknown length
8211: gap of unknown length
8310: gap of unknown length
8311: gap of unknown length
9467: contig of 1157 bp in length
9567: gap of unknown length
10766: contig of 1199 bp in length
10767: gap of unknown length
10866: gap of unknown length
11994: contig of 1128 bp in length
12034: gap of unknown length
12095: gap of unknown length
13102: contig of 1008 bp in length
13103: gap of unknown length
13202: gap of unknown length
14497: contig of 1295 bp in length
14597: gap of unknown length
14598: contig of 1541 bp in length
16138: gap of unknown length
16238: gap of unknown length
17906: contig of 1668 bp in length
18006: gap of unknown length
18007: gap of unknown length
19573: contig of 1567 bp in length
19574: gap of unknown length
19575: gap of unknown length
20704: contig of 1031 bp in length
20804: gap of unknown length
22301: contig of 1497 bp in length
22401: gap of unknown length
24843: contig of 2442 bp in length
24844: gap of unknown length
26418: contig of 1475 bp in length
26419: gap of unknown length
26518: gap of unknown length
28360: contig of 1842 bp in length
28361: gap of unknown length
30423: contig of 1963 bp in length
30523: gap of unknown length
30524: gap of unknown length
32559: contig of 2036 bp in length
32659: gap of unknown length
32660: gap of unknown length
34631: contig of 1972 bp in length
34632: gap of unknown length
34731: gap of unknown length
37171: contig of 2440 bp in length
37172: gap of unknown length
37271: gap of unknown length
39423: contig of 2152 bp in length
39523: gap of unknown length
41124: contig of 1601 bp in length
39524

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 23, 2003, 04:38:33 ; Search time 366 Seconds
(without alignments)
9555.610 Million cell updates/sec

Title: US-09-865-879-2

Perfect score: 1553
Sequence: 1 gactcagggtgtcccaaac.....ggtcgcgtcgtccgcctcc 1553

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq.101002:*

1: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*

2: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*

3: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*

4: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*

5: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*

6: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*

7: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*

8: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*

9: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*

10: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*

11: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*

12: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*

13: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*

14: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*

15: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*

16: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*

17: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*

18: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*

19: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*

20: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*

21: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*

22: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*

23: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*

24: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1553	100.0	1553	24	AAD24913
2	59	3.8	2790	24	AAS94998
3	51.4	3.3	883	22	AAU26741
4	50.4	3.2	3037	22	AAK94933
5	49.4	3.2	7998	22	AAK79404
6	47.6	3.1	407	21	AAK7895
7	47.6	3.1	407	22	AAI28633
8	47.4	3.1	114955	20	AAK53491
9	47	3.0	2691	14	AAO48043

10	47	3.0	2691	16	AAO98815
11	47	3.0	2691	21	AAA7842
12	47	3.0	2691	22	AAI28580
13	47	3.0	2691	24	ABN95902
14	47	3.0	2691	24	ABR62672
15	46.6	3.0	1907	24	ABR99476
16	46.2	3.0	174424	24	ABL68122
17	46	3.0	53522	24	AAD30228
18	46	3.0	53526	19	AAR94101
19	46	3.0	53577	17	AAT18551
20	46	3.0	53577	19	AAR94108
21	46	3.0	105325	24	ABR94407
22	45.6	2.9	368	22	AAH09431
23	45.4	2.9	1166	20	AAI19440
24	45.4	2.9	1166	20	AAI19428
25	45	2.9	7620	22	ABA07349
26	45	2.9	7620	22	AAK32776
27	45	2.9	114955	20	AAK53491
28	44.4	2.9	1154	24	ABO68318
29	44.4	2.9	1523	20	AAI17507
30	44.4	2.9	1523	20	AAI17479
31	44.4	2.9	1608	21	AAI22298
32	44.2	2.8	1050	22	AAD21685
33	44.2	2.8	2849	22	AAD21684
34	43.6	2.8	357	22	AAI181967
35	43.4	2.8	220895	24	ABR84798
36	43.4	2.8	117213	19	AAV62176
37	43.2	2.8	169739	24	ABO88186
38	42.8	2.8	3441	20	AAK80044
39	42.8	2.8	3441	21	AAA49552
40	42.8	2.8	3441	24	ABR40258
41	42.8	2.8	3441	24	ABR28570
42	42.6	2.7	600	24	ABO52496
43	42.6	2.7	600	24	ABO52497
44	42.6	2.7	100301	24	ABO88176
45	42.4	2.7	1362	15	AAO64890

ALIGNMENTS

RESULT 1
ID AAD24913 standard; DNA; 1553 BP.
XX AAD24913:
XX
XX 12-MAR-2002 (first entry)
XX
XX Human secreted cell adhesion protein beta IG-H3 promoter DNA.
XX
XX Human: growth inhibitory gene; retinoid; retinoic acid response element;
XX RARE site; therapy; promyelocytic leukaemia; cancer chemoprevention;
XX cytosolic; secreted cell adhesion protein beta IG-H3 promoter; ds.
XX
XX Homo sapiens.
XX
XX WO200192578-A2.
XX
XX 06-DEC-2001.
XX
XX 25-MAY-2001; 2001WO-US17161.
XX
XX 26-MAY-2000; 2000US-207535P.
XX
XX (UNIT) UNIV ILLINOIS FOUNO.
XX
XX Roninson IB, Dokmanovic M, Chang B;
XX WPI, 2002-075474/10.
XX
XX Expression construct encoding cellular genes, under control of a
XX promoter regulated by retinoids and cells comprising the construct for
XX PT

PT Identifying compounds that induce expression of the genes useful in
PT treating cancer -
XX
PS Claim 4, Page 43-44; 64pp. English.

CC The patent discloses growth inhibitory genes induced by retinoids. The
CC invention also relates to recombinant expression constructs that express
CC a reporter gene under the transcriptional control of a promoter for a
CC gene which is expressed by retinoid induction. The promoter does not
CC contain a retinoic acid response elements (RARE) site. The invention
CC further relates to reagents and methods for identifying compounds other
CC than retinoids that modulate the expression of cellular genes. These
CC compounds are useful for treating cancers such as promyelocytic leukaemia
CC and cancer chemoprevention. The present sequence is human secreted cell
CC adhesion protein beta Ig-H3 promoter DNA used in the invention.

Sequence 1553 BP; 305 A; 490 C; 384 G; 374 T; 0 other;

Query Match	100.0%	Score 1553;	DB 24;	Length 1553;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1553;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY	1	GACTCAGGGGTGCCAAACACACACTATCACCCTGGCAACCCCTGCACCTCGCATGTCGCTCA	60
Db	1	GACTCAGGGGTGCCAAACACACACTATCACCCTGGCAACCCCTGCACCTCGCATGTCGCTCA	60
QY	61	TTTCGAAACATGGCACACATCATCTGCTGCAATGTCACAGACACAAACACCCCTACAAATATCT	120
Db	61	TTTCGAAACATGGCACACATCATCTGCTGCAATGTCACAGACACAAACACCCCTACAAATATCT	120
QY	121	TGACCTCTCTTCGCCCTTCTCCCTGTATACAGACGTCACAAATTCATATGAGACTATTAC	180
Db	121	TGACTCTCTTCTTCGCCCTTCTCCCTGTATACAGACGTCACAAATTCATATGAGACTATTAC	180
QY	181	CTCTACACCCCTGACATTTGTCACAGCCCTTCCCATCTCTGACCTCTACACACATAGTTCA	240
Db	181	CTCTACACCCCTGACATTTGTCACAGCCCTTCCCATCTCTGACCTCTACACACATAGTTCA	240
QY	241	AGCTCTCCCATGATGTCCTTCTCTGTTACCTCTTCTTCTGCTCTCTTAAGCTCTCATGA	300
Db	241	AGCTCTCCCATGATGTCCTTCTCTGTTACCTCTTCTTCTGCTCTCTTAAGCTCTCATGA	300
QY	301	CACATGGCCATTCACATTCCTCTCCACCATACACCCGGATAGGCTTACTCTGAGCTGCGG	360
Db	301	CACATGGCCATTCACATTCCTCTCCACCATACACCCGGATAGGCTTACTCTGAGCTGCGG	360
QY	361	CCCTGCTACCTTCCTCCCTTCCTTCCTTACCCTTACCTACCTCCCTGCTTACGCCAA	420
Db	361	CCCTGCTACCTTCCTCCCTTCCTTCCTTACCCTTACCTACCTCCCTGCTTACGCCAA	420
QY	421	CCAGATTAACCTTGAGCTTTCGTGTAATGATGACCTCACTTACCTGATTAACATCATCTATTC	480
Db	421	CCAGATTAACCTTGAGCTTTCGTGTAATGATGACCTCACTTACCTGATTAACATCATCTATTC	480
QY	481	TTTCAGGCTCAGAGGAGGTATCACCCCTGTACAGGCGCAGGCGCTCTTCTTAGCTCCAA	540
Db	481	TTTCAGGCTCAGAGGAGGTATCACCCCTGTACAGGCGCAGGCGCTCTTCTTAGCTCCAA	540
QY	541	AGCCCCAGCTACTTTCATGAGAAATCATTTGGCTTGGGCTACGATCTTCCCAATTGGA	600
Db	541	AGCCCCAGCTACTTTCATGAGAAATCATTTGGCTTGGGCTACGATCTTCCCAATTGGA	600
QY	601	GCCTTTTACAAAGGGCTTAGGTCACATCTATTAATTCATCTGTGTCCTCCACG	660
Db	601	GCCTTTTACAAAGGGCTTAGGTCACATCTATTAATTCATCTGTGTCCTCCACG	660
QY	661	GCTTAGCAGTGCAGTAATCTGACAGGATTAATAGATGCTTGGGTAAGTATCACCTCT	720
Db	661	GCTTAGCAGTGCAGTAATCTGACAGGATTAATAGATGCTTGGGTAAGTATCACCTCT	720
QY	721	TACCATGAGCAATTTGTTTATACCGCCCTTAGGCTCCACAGGACAGACATCTTGCTTG	780
Db	721	TACCATGAGCAATTTGTTTATACCGCCCTTAGGCTCCACAGGACAGACATCTTGCTTG	780
QY	780	TACCATGAGCAATTTGTTTATACCGCCCTTAGGCTCCACAGGACAGACATCTTGCTTG	780
Db	780	TACCATGAGCAATTTGTTTATACCGCCCTTAGGCTCCACAGGACAGACATCTTGCTTG	780

Chr	Start (kb)	End (kb)	Gene	Accession	Length (bp)	GC (%)	GC3 (%)	GC4 (%)	GC5 (%)	GC6 (%)	GC7 (%)	GC8 (%)	GC9 (%)	GC10 (%)	GC11 (%)	GC12 (%)	GC13 (%)	GC14 (%)	GC15 (%)	GC16 (%)	GC17 (%)	GC18 (%)	GC19 (%)	GC20 (%)	GC21 (%)	GC22 (%)	GC23 (%)	GC24 (%)	GC25 (%)	GC26 (%)	GC27 (%)	GC28 (%)	GC29 (%)	GC30 (%)	GC31 (%)	GC32 (%)	GC33 (%)	GC34 (%)	GC35 (%)	GC36 (%)	GC37 (%)	GC38 (%)	GC39 (%)	GC40 (%)	GC41 (%)	GC42 (%)	GC43 (%)	GC44 (%)	GC45 (%)	GC46 (%)	GC47 (%)	GC48 (%)	GC49 (%)	GC50 (%)	GC51 (%)	GC52 (%)	GC53 (%)	GC54 (%)	GC55 (%)	GC56 (%)	GC57 (%)	GC58 (%)	GC59 (%)	GC60 (%)	GC61 (%)	GC62 (%)	GC63 (%)	GC64 (%)	GC65 (%)	GC66 (%)	GC67 (%)	GC68 (%)	GC69 (%)	GC70 (%)	GC71 (%)	GC72 (%)	GC73 (%)	GC74 (%)	GC75 (%)	GC76 (%)	GC77 (%)	GC78 (%)	GC79 (%)	GC80 (%)	GC81 (%)	GC82 (%)	GC83 (%)	GC84 (%)	GC85 (%)	GC86 (%)	GC87 (%)	GC88 (%)	GC89 (%)	GC90 (%)	GC91 (%)	GC92 (%)	GC93 (%)	GC94 (%)	GC95 (%)	GC96 (%)	GC97 (%)	GC98 (%)	GC99 (%)	GC100 (%)	GC101 (%)	GC102 (%)	GC103 (%)	GC104 (%)	GC105 (%)	GC106 (%)	GC107 (%)	GC108 (%)	GC109 (%)	GC110 (%)	GC111 (%)	GC112 (%)	GC113 (%)	GC114 (%)	GC115 (%)	GC116 (%)	GC117 (%)	GC118 (%)	GC119 (%)	GC120 (%)	GC121 (%)	GC122 (%)	GC123 (%)	GC124 (%)	GC125 (%)	GC126 (%)	GC127 (%)	GC128 (%)	GC129 (%)	GC130 (%)	GC131 (%)	GC132 (%)	GC133 (%)	GC134 (%)	GC135 (%)	GC136 (%)	GC137 (%)	GC138 (%)	GC139 (%)	GC140 (%)	GC141 (%)	GC142 (%)	GC143 (%)	GC144 (%)	GC145 (%)	GC146 (%)	GC147 (%)	GC148 (%)	GC149 (%)	GC150 (%)	GC151 (%)	GC152 (%)	GC153 (%)	GC154 (%)	GC155 (%)	GC156 (%)	GC157 (%)	GC158 (%)	GC159 (%)	GC160 (%)	GC161 (%)	GC162 (%)	GC163 (%)	GC164 (%)	GC165 (%)	GC166 (%)	GC167 (%)	GC168 (%)	GC169 (%)	GC170 (%)	GC171 (%)	GC172 (%)	GC173 (%)	GC174 (%)	GC175 (%)	GC176 (%)	GC177 (%)	GC178 (%)	GC179 (%)	GC180 (%)	GC181 (%)	GC182 (%)	GC183 (%)	GC184 (%)	GC185 (%)	GC186 (%)	GC187 (%)	GC188 (%)	GC189 (%)	GC190 (%)	GC191 (%)	GC192 (%)	GC193 (%)	GC194 (%)	GC195 (%)	GC196 (%)	GC197 (%)	GC198 (%)	GC199 (%)	GC200 (%)	GC201 (%)	GC202 (%)	GC203 (%)	GC204 (%)	GC205 (%)	GC206 (%)	GC207 (%)	GC208 (%)	GC209 (%)	GC210 (%)	GC211 (%)	GC212 (%)	GC213 (%)	GC214 (%)	GC215 (%)	GC216 (%)	GC217 (%)	GC218 (%)	GC219 (%)	GC220 (%)	GC221 (%)	GC222 (%)	GC223 (%)	GC224 (%)	GC225 (%)	GC226 (%)	GC227 (%)	GC228 (%)	GC229 (%)	GC230 (%)	GC231 (%)	GC232 (%)	GC233 (%)	GC234 (%)	GC235 (%)	GC236 (%)	GC237 (%)	GC238 (%)	GC239 (%)	GC240 (%)	GC241 (%)	GC242 (%)	GC243 (%)	GC244 (%)	GC245 (%)	GC246 (%)	GC247 (%)	GC248 (%)	GC249 (%)	GC250 (%)	GC251 (%)	GC252 (%)	GC253 (%)	GC254 (%)	GC255 (%)	GC256 (%)	GC257 (%)	GC258 (%)	GC259 (%)	GC260 (%)	GC261 (%)	GC262 (%)	GC263 (%)	GC264 (%)	GC265 (%)	GC266 (%)	GC267 (%)	GC268 (%)	GC269 (%)	GC270 (%)	GC271 (%)	GC272 (%)	GC273 (%)	GC274 (%)	GC275 (%)	GC276 (%)	GC277 (%)	GC278 (%)	GC279 (%)	GC280 (%)	GC281 (%)	GC282 (%)	GC283 (%)	GC284 (%)	GC285 (%)	GC286 (%)	GC287 (%)	GC288 (%)	GC289 (%)	GC290 (%)	GC291 (%)	GC292 (%)	GC293 (%)	GC294 (%)	GC295 (%)	GC296 (%)	GC297 (%)	GC298 (%)	GC299 (%)	GC300 (%)	GC301 (%)	GC302 (%)	GC303 (%)	GC304 (%)	GC305 (%)	GC306 (%)	GC307 (%)	GC308 (%)	GC309 (%)	GC310 (%)	GC311 (%)	GC312 (%)	GC313 (%)	GC314 (%)	GC315 (%)	GC316 (%)	GC317 (%)	GC318 (%)	GC319 (%)	GC320 (%)	GC321 (%)	GC322 (%)	GC323 (%)	GC324 (%)	GC325 (%)	GC326 (%)	GC327 (%)	GC328 (%)	GC329 (%)	GC330 (%)	GC331 (%)	GC332 (%)	GC333 (%)	GC334 (%)	GC335 (%)	GC336 (%)	GC337 (%)	GC338 (%)	GC339 (%)	GC340 (%)	GC341 (%)	GC342 (%)	GC343 (%)	GC344 (%)	GC345 (%)	GC346 (%)	GC347 (%)	GC348 (%)	GC349 (%)	GC350 (%)	GC351 (%)	GC352 (%)	GC353 (%)	GC354 (%)	GC355 (%)	GC356 (%)	GC357 (%)	GC358 (%)	GC359 (%)	GC360 (%)	GC361 (%)	GC362 (%)	GC363 (%)	GC364 (%)	GC365 (%)	GC366 (%)	GC367 (%)	GC368 (%)	GC369 (%)	GC370 (%)	GC371 (%)	GC372 (%)	GC373 (%)	GC374 (%)	GC375 (%)	GC376 (%)	GC377 (%)	GC378 (%)	GC379 (%)	GC380 (%)	GC381 (%)	GC382 (%)	GC383 (%)	GC384 (%)	GC385 (%)	GC386 (%)	GC387 (%)	GC388 (%)	GC389 (%)	GC390 (%)	GC391 (%)	GC392 (%)	GC393 (%)	GC394 (%)	GC395 (%)	GC396 (%)	GC397 (%)	GC398 (%)	GC399 (%)	GC400 (%)	GC401 (%)	GC402 (%)	GC403 (%)	GC404 (%)	GC405 (%)	GC406 (%)	GC407 (%)	GC408 (%)	GC409 (%)	GC410 (%)	GC411 (%)	GC412 (%)	GC413 (%)	GC414 (%)	GC415 (%)	GC416 (%)</
-----	------------	----------	------	-----------	-------------	--------	---------	---------	---------	---------	---------	---------	---------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-------------

```

PF 04-APR-2001; 2001WO-US11128.
XX
PR 05-APR-2000; 2000US-195106P.
PR
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Shiffman D, Somogyi R, Lawn R, Sellhammer JJ, Porter GJ, Mikita T;
PI Tal J;
XX
XX WPI; 2002-010925/01.
XX
PT Composition useful for diagnosis of conditions, disorders or diseases
PT associated with atherosclerosis, comprises several polynucleotides that
PT are differentially expressed in foam cell development -
XX
PS Claim 1; Page 295-196; 315pp; English.
XX
XX The present invention relates to the isolation of human polynucleotide
CC sequences that are differentially expressed during foam cell
CC differentiation. The polynucleotide sequences of the invention or a
CC composition comprising these polynucleotides are useful as a high
CC throughput method for detecting altered expression of one or more
CC polynucleotides in a sample. The polynucleotides can be used in the
CC diagnosis of disorders associated with foam cell development such as
CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as
CC coronary artery disease. The polynucleotide sequences can also be used
CC as PCR primers and probes. The polynucleotides of the invention are also
CC useful in gene therapy. AAS94746-AAS95021 represent the human
CC polynucleotide sequences of the invention which are differentially
CC expressed during foam cell differentiation.
XX
SQ Sequence 2790 BP; 708 A; 756 C; 716 G; 610 T; 0 other;
XX
XX Query Match 3.8%; Score 59; DB 24; Length 2790;
XX Best Local Similarity 100.0%; Pred. No. 2.5e-05;
XX Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0
XX
OY 1495 CTGAGAGCCGCCGCTTGCCCGTGGTCGCTGACTGCTGCGGTGCGGTCCGCTCC 1553
XX |||||||
XX 1 CTGAGAGCCGCCGCTTGCCCGTGGTCGCTGACTGCTGCGGTGCGGTCCGCTCC 59
XX
RESULT 3
AAL26741
ID AAL26741 standard; cDNA; 883 BP.
XX
XX AAL26741;
XX
XX 07-DEC-2001 (first entry)
XX
XX Human breast cancer expressed polynucleotide 19198.
XX
XX Human; breast cancer; cell marker; cytostatic; ss.
XX
XX Homo sapiens.
XX
XX W0200151628-A2.
XX
XX 19-JUL-2001.
XX
XX 10-JAN-2001; 2001WO-US00798.
XX
XX 14-JAN-2000; 2000US-0176077.
XX 14-MAR-2000; 2000US-0189167.
XX 24-MAR-2000; 2000US-0192099.
XX 29-MAR-2000; 2000US-0193480.
XX 15-MAY-2000; 2000US-0205230.
XX 09-JUN-2000; 2000US-0211315.
XX 25-JUL-2000; 2000US-0220534.
XX
XX (MILT-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX

```

```

XX WP1: 2001-451856/48.
DR
XX
XX New peptide useful as a marker for the diagnosis of breast cancer
PT
XX
XX Claim 1; Page 3668-3669; 3695pp; English.
PS
XX
XX The invention relates to human breast cancer expressed polynucleotides
CC CC
CC (AA107544-AA126789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterizing treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity.
XX
XX
SQ Sequence 883 BP; 175 A; 274 C; 251 G; 183 T; 0 other;
Query Match 3.3%; Score 51.4; DB 22; Length 883;
Best Local Similarity 90.2%; Pred. No. 0.0016;
Matches 55; Conservative 0; Mismatches 6; Indels 0; Gaps
OY 1493 CCCGTGAGACCGCCGCTGCGCGTGGCGTACGTGCTGCGTGGCGTCCGCTC 1552
Db 13 CCCACGCGTCCGCGCGCTGCGCGTGGCGTACGTGCTGCGTGGCGTCCGCTC 72
OY 1553 C 1553
Db 73 C 73
RESULT 4
ID AAK94933/C AAK94933 standard; cDNA; 3037 BP.
AC AAK94933;
AAK94933;
DT 06-NOV-2001 (first entry)
DE Human full-length cDNA, SEQ ID NO: 4176.
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
XX Homo sapiens.
OS
XX
XX EP1130094-A2.
XX
XX 05-SEP-2001.
XX
XX 07-JUL-2000; 2000EP-0114089.
XX
XX 08-JUL-1999; 99JP-0194486.
XX 11-JAN-2000; 2000JP-0118774.
XX 02-MAY-2000; 2000JP-0183765.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI: 2001-524255/58.
XX P-PSDB; AAM93965.
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their
XX use in genetic manipulation -
XX
XX Claim 8; SEQ ID NO 4176; 1380bp + sequence listing; English.
XX
XX The invention relates to primers for synthesizing full length cDNA
XX clones. 830 cDNA molecules encoding a human protein have been
XX isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
XX molecules have been determined. Primers for synthesizing the full length

```


08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251889.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX
PA (HUMA-) HUMAN GENOME SCT INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
DR
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
XX
PS Disclosure; SEQ ID NO 34216; 3071bp + Sequence Listing; English.
XX
XX
CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially

CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 7998 BP; 1435 A; 2527 C; 2461 G; 1575 T; 0 other;
Query Match 3.2%; Score 49.4; DB 22; Length 7998;
Best Local Similarity 57.1%; Pred. No. 0.014;
Matches 109; Conservative 0; Mismatches 81; Indels 1; Gaps 1;
QY 1267 GAGAGGAGGAGGAGGCTGCGAGTGAAGGCAAGGCTGGGAAAAGTGAAGCAGCGGACAGT 1326
DB 1609 GCGGGAGGAGGAGGAGCGCGGAGTGGGGGCGCGGC-GTGGCTCTGCTCCGGAGCGCGGC 1551
QY 1327 GCGGAGCGGGGTGGTCCAGGAGCCAGCGGAGGCGACAGGGTGGAGGCGCAGCGCGG 1386
DB 1550 TCTGCTGCAAGTGCAGATCTTGAAGCCAGGAGGCGGCGTGGGAGGCGGAGTCAAG 1491
QY 1387 CCGGCCCTCTTGCACAGGCGGCGCCAGCTTCCCGCCCTGGCGTCCGCTCCCTCCCGC 1446
DB 1490 CTCCTTGACTTGGGTGGGCGCCCTCTCGCCTCTCCGCTTACGCAAGCGCTCTACTCTGA 1431
QY 1447 TCGCAGCTTAC 1457
DB 1430 GCGAGCGCTAC 1420
RESULT 6
ID AAK77895
ID AAK77895 standard; cDNA; 407 BP.
XX
XX AAK77895;
AC
XX
DT 14-NOV-2000 (first entry)
XX
XX cDNA encoding human colon tumour polypeptide, SEQ ID NO:175.
DE
XX
XX Human colon tumour polypeptide; tumour antigen; cancer; vaccine;
KW immunotherapy; diagnosis; progression; ss.
XX
XX Homo sapiens.
PN WO200037643-A2.
XX
XX 29-JUN-2000.
PD
XX
XX 23-DEC-1999; 99WO-US30909.
PF
XX
XX 23-DEC-1998; 98US-0221298.
PR 02-JUL-1999; 99US-0347496.
PR 22-SEP-1999; 99US-0401064.
PR 19-NOV-1999; 99US-0444242.
PR 02-DEC-1999; 99US-0454150.
XX
XX (CORI-) CORIXA CORP.
XX
XX
PI Xu J, Lodes MJ, Secret H, Benson DR, Meagher MJ, Stolk J;
PI Wang T, Yugu J;
XX
XX WPI; 2000-442671/38.
DR
XX
XX New colon tumor polypeptides used to inhibit the development of cancer,
PT especially colon cancer, and for diagnosing and monitoring the
PT progression of the cancer -
XX
XX
PS Claim 1; Page 142; 229pp; English.
XX
XX Sequences AAK77722-A78199 represent 478 cDNAs encoding proteins or
CC portions of proteins which are associated with human colon tumours.
CC The invention also specifically discloses 8 human colon tumour proteins
CC (AA11897-B11904). The nucleic acids, the polypeptides they encode, and
CC antigen presenting cells (APCs, preferably dendritic cells) expressing

XX	RESULT	10
XX	AAQ98815	
XX	ID	AAQ98815 standard; DNA; 2691 BP.
XX	AC	
XX	AAQ98815;	
XX	DPT	06-MAR-1996 (first entry)
XX	DE	Human beta-IG-H3 (transforming growth factor-beta induced gene-h3).
XX	KW	beta-IG-H3; antitumor; ss.
OS	Homo sapiens.	
XX		
FT	key	location/Qualifiers
FT	CDS	48..2099
FT		/tag= a
FT	sig_peptide	48..89
FT		/tag= b
FT	misc_feature	89..90
FT		/tag= c
FT	misc_feature	/note= "cleavage site"
FT		116..117
FT		/tag= d
FT	polyA_signal	/note= "cleavage site"
FT		2669..2677
FT		/tag= e
FT	repeat_region	462..869
FT		/tag= f
FT	repeat_region	/note= "REPEAT 1"
FT		870..1277
FT	repeat_region	/note= "REPEAT 2"
FT		1278..1658
FT	repeat_region	/tag= h
FT		/note= "REPEAT 3"
FT		1659..2693
FT		/tag= i
FT		/note= "REPEAT 4"
XX	US5444164-A.	
XX	PD	22-AUG-1995.
XX	PE	05-FEB-1992; 92US-0833835.
XX	PR	04-MAY-1992; 92US-0878960.
XX	PR	05-FEB-1992; 92US-0833835.
PA	(BRIM) BRISTOL-MYERS SQUIBB CO.	
PI	Neubauer MG, Purchio AF, Skonier JE;	
DR	WPI; 1995-302161/39.	
DR	P-PSDB; AAR80573.	
XX		
PT	New isolated nucleic acid and protein, beta IG-H3 - which are	
PT	induced by mammalian cells treated with TGF-beta, used for	
PT	inhibiting growth of tumor cells	
PS	Claim 1; Page 11-14; 23pp; English.	
CC	DNA encoding betaIG-H3 was isolated from a cDNA library prepared	
CC	from A549 human lung adenocarcinoma cells treated with transforming	
CC	growth factor-beta. This sequence encodes a beta-IG-H3 protein, th	
CC	induction, expression and/or secretion of which inhibits the growth	
CC	colony formation and appearance of tumor cells. The beta-IG-H3 can	
CC	be used to inhibit the growth of tumor cells.	
Sequence	2691 BP; 679 A; 729 C; 695 G; 588 T; 0 other;	

Query Match: 3.0%; Score 47; DB 16; Length 2691;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

Dy 1507 GCCTGCCCGTCGTCGTACTACGTCGTCGGTCCGTCGTCGCATC 1553
|||||
Db 1 GCTTGCCCGTCGTCGTACTACGTCGTCGGTCCGTCGTCGCATC 47

RESULT 11
AAA77842
ID AAA77842 standard; cDNA; 2691 BP.
XX
AC AAA77842:
XX
DT 14-NOV-2000 (first entry)
XX
DE cDNA encoding human colon tumour polypeptide, SEQ ID NO:121.
XX
KW Human colon tumour polypeptide; tumour antigen; cancer; vaccine;
KW immunotherapy; diagnosis; progression; beta IG-H3 homologue; ss.
XX
OS Homo sapiens.
XX
PN MO200037643-A2.
PD 29-JUN-2000.
XX
PE 23-DEC-1999; 99WO-US30909.
XX
PR 23-DEC-1998; 98US-0221298.
PR 02-JUL-1999; 99US-0347496.
PR 22-SEP-1999; 99US-0401064.
PR 19-NOV-1999; 99US-0444242.
XX 02-DEC-1999; 99US-0454150.
PA (CORI-) CORIXA CORP.
PI Xu J, Lodes MJ, Secretist H, Benson DR, Meagher MJ, Stolk J;
PI Wang T, Yugu J;
PT
PT
PS P-PSDB; AAB11897.
DR WPI: 2000-442671/38.
DR
XX
XX
XX New colon tumor polypeptides used to inhibit the development of cancer,
especially colon cancer, and for diagnosing and monitoring the
progression of the cancer -
Claim 29; Page 123; 229pp; English.

Sequences AAA77722-A78199 represent 478 cDNAs encoding proteins or
portions of proteins which are associated with human colon tumours.
The invention also specifically discloses 8 human colon tumour proteins
(AAB11897-B11904). The nucleic acids, the polypeptides they encode, and
antigen presenting cells (APCs, preferably dendritic cells) expressing
such polypeptides may be used in vaccines that target tumour cells,
especially colon tumour cells, thereby inhibiting the development of
cancer. T-cells specific for the polypeptide expressed by the APC are
used to remove tumour cells from biological samples, especially blood or
fractions thereof. The sample or the isolated T-cells specific for the
polypeptide can then be used to inhibit cancer development. CD4+ and/or
CD8+ T-cells from a patient may be incubated with a polypeptide or
nucleic acid of the invention, or an APC expressing such a polypeptide,
to cause the proliferation of specific T-cells. The T-cells can be
cloned and then administered back to the patient to inhibit cancer
development. Nucleic acids encoding the polypeptides and antibodies
against the polypeptides may be used to determine the expression level
of a tumour protein of the invention, and therefore to determine whether
cancer cells are present. Such diagnostic methods may also be used to
monitor the progression of a cancer by repeating the processes at time
intervals, and comparing the current result to previous results. The
present sequence represents a cDNA encoding a human colon tumour
polypeptide which has homology to beta IG-H3.

XX	Sequence	2691 BP;	679 A;	730 C;	695 G;	587 T;	0 other;
SO	Query Match	3.0%;	Score 47;	DB 21;	Length 2691;		
	Best Local Similarity	100.0%;	Pred. No.	0.039;			
	Matches 47;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		
OY	1507	GCCTGCCCCGTGGTGGTCAGCGCCGTCGTCGCCGCGCCGCCGCC	1553				
DB	1	GCCTGCCCCGTGGTGGTCAGCTCCTCGTGCGCTGCGCTGCTCCGCTCC	47				
	RESULT 12						
ID	AAI28580						
XX	AAI28580 standard; cDNA; 2691 BP.						
AC	AAI28580;						
XX							
DT	12-OCT-2001 (first entry)						
XX							
DE	Colon tumour related full length cDNA sequence SEQ ID NO:121.						
XX							
KW	Human; immunotherapy; diagnosis; colon cancer; colon tumour;						
KW	immunogenic; gene therapy; vaccine; colonic cancer; ss.						
XX							
OS	Homo sapiens.						
XX							
PN	WO200149716-A2.						
PD	12-JUL-2001.						
XX							
PF	29-DEC-2000; 2000WO-US35596.						
XX							
PR	30-DEC-1999; 99US-0476296.						
PR	10-JAN-2000; 2000US-0480321.						
PR	15-FEB-2000; 2000US-0504629.						
PR	06-MAR-2000; 2000US-0519444.						
PR	19-MAY-2000; 2000US-0575251.						
PR	29-JUN-2000; 2000US-0609448.						
PR	28-AUG-2000; 2000US-0649811.						
XX							
PA	(CORI-) CORIXA CORP.						
XX							
PI	Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;						
PI	King GE, Wang T, Jiang Y;						
XX							
DR	WPL; 2001-441847/47.						
XX							
PT	Colon tumor associated proteins and nucleic acids useful for the						
PT	prevention, diagnosis and treatment of colonic cancer -						
PS	claim 25; Page 164-165; 472pp; English.						
XX							
CC	The present invention describes colon tumour associated proteins (I) and						
CC	the polynucleotides (II) that encode them. (I) have cytostatic activity.						
CC	(I) and (II) can be used in gene therapy and vaccine production. (I) and						
CC	(II) may be used in the prevention, diagnosis and treatment of diseases						
CC	associated with inappropriate colon tumour associated protein (TCAP)						
CC	expression, such as colonic cancer. For example, (I) and (II) may be						
CC	used to treat disorders associated with decreased expression by						
CC	rectifying mutations or deletions in a patient's genome that affect the						
CC	activity of TCAPs by expressing inactive proteins or to supplement the						
CC	patient's own production of them. Additionally, (II) may be used to						
CC	produce the TCAP proteins, by inserting the nucleic acids into a host						
CC	cell culturing the cell to express the protein. (II) and its						
CC	complementary sequences may also be used as DNA probes in diagnostic						
CC	polymerase chain reaction (PCR) and hybridisation assays to detect and						
CC	quantitate the presence of similar nucleic acids in samples, and						
CC	therefore which patients may be in need of restorative therapy. (I) may						
CC	also be used as antigens in the production of antibodies against TCAPs						
CC	and in assays to identify modulators of TCAP expression and activity.						
CC	Anti-(I) antibodies and antagonists may also be used to down regulate						
CC	TCAP expression and activity. The anti-(I) antibodies may also be used						

CC	as diagnostic agents for detecting the presence of TCAPs in samples
CC	(e.g., by enzyme linked immunosorbent assay (ELISA)). AA128460 to AA129512
CC	and AAM24494 to AAM24523 represent nucleotide and amino acid sequences
CC	given in the exemplification of the present invention.
XX	
SQ	Sequence 2691 BP; 679 A; 729 C; 695 G; 588 T; 0 other;
Query Match	3.0%; Score 47; DB 22; Length 2691;
Best Local Similarity	100.0%; Pred. No. 0.039;
Matches 47; Conservative	0; Mismatches 0; Indels 0; Gaps 0
OY	1507 GCCTGCCCGTCGTCGTAGCTCCTGGTGGCGGCATCCCCGTTCC 1553
Db	1 GCTTGCCCGTCGTCGTAGCTCCTGGTGGCGGCATCCCCGTTCC 47
RESULT 13	
ID	ABN95902
ABN95902 standard; DNA; 2691 BP.	
XX	
AC	ABN95902:
XX	
DT	13-AUG-2002 (first entry)
XX	
DE	Gene #2400 used to diagnose liver cancer.
XX	
KW	Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic; metastatic liver tumour; cytostatic; expression profile; disease state; disease progression; drug toxicity; drug efficacy; drug metabolism.
XX	
OS	Homo sapiens.
XX	
PN	WO200229103-A2.
PD	11-APR-2002.
PE	02-OCT-2001; 2001WO-US30589.
PR	02-OCT-2000; 2000US-237054P.
PA	(GENE-) GENE LOGIC INC.
XX	
PI	Horne D, Alvares C, Peres-Da-Silva S, Vockley JC;
DR	WPI; 2002-426119/45.
PT	Diagnosing and detecting the progression of liver cancer,
PT	hepatocellular carcinoma or metastatic liver tumor in a patient,
PT	involves detecting the level of expression of two or more genes in a
XX	liver tissue sample -
PS	Claim 1; SEQ ID NO 2400; 298pp; English.
XX	
CC	The invention relates to a novel method for diagnosing and detecting the
CC	progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC	tumour in a patient, and differentiating metastatic liver cancer from
CC	hepatocellular carcinoma in a patient, involving detecting the level of
CC	expression of two or more genes represented in ABN93503-ABN97455 in a
CC	tissue sample. The method of the invention has hepatotropic, and
CC	cytostatic activity. The method is useful for diagnosing and detecting
CC	the progression of liver cancer, hepatocellular carcinoma and metastatic
CC	liver carcinoma in a patient. The method is useful for identifying
CC	expression profiles which serve as useful diagnostic markers as well as
CC	markers that can be used to monitor disease states, disease progression,
CC	drug toxicity, drug efficacy and drug metabolism.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	
SQ	Sequence 2691 BP; 679 A; 729 C; 695 G; 588 T; 0 other;
Query Match	3.0%; Score 47; DB 24; Length 2691;
Best Local Similarity	100.0%; Pred No. 0.039;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1507 GCTGGCCGTCGCTAGCTGCTGCGGCGTCGCCGCC 1553
 |||||||
 Db 1 GCTGGCCGTCGCTAGCTGCTGCGGCGTCGCCGCC 47

RESULT 14
 ABL62672
 ID ABL62672 standard; DNA: 2691 BP.

AC ABL62672;

DT 15-MAY-2002 (first entry)

DE Colon adenocarcinoma related gene sequence SEQ ID NO:1009.

KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 cytosolic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 gene; ds.

XX Homo sapiens.

PN WO200194629-A2.

PD 13-DEC-2001.

PF 30-MAY-2001; 2001WO-US10838.

PR 05-JUN-2000; 2000US-209473P.

PR 05-JUN-2000; 2000US-209531P.

PR 18-SEP-2000; 2000US-233133P.

PR 18-SEP-2000; 2000US-233617P.

PR 20-SEP-2000; 2000US-234009P.

PR 20-SEP-2000; 2000US-234034P.

PR 20-SEP-2000; 2000US-234052P.

PR 22-SEP-2000; 2000US-234509P.

PR 22-SEP-2000; 2000US-234567P.

PR 25-SEP-2000; 2000US-234923P.

PR 25-SEP-2000; 2000US-234924P.

PR 25-SEP-2000; 2000US-235082P.

PR 25-SEP-2000; 2000US-235082P.

PR 25-SEP-2000; 2000US-235134P.

PR 25-SEP-2000; 2000US-235280P.

PR 26-SEP-2000; 2000US-235637P.

PR 26-SEP-2000; 2000US-235638P.

PR 27-SEP-2000; 2000US-235711P.

PR 27-SEP-2000; 2000US-235720P.

PR 27-SEP-2000; 2000US-235840P.

PR 27-SEP-2000; 2000US-235863P.

PR 28-SEP-2000; 2000US-236028P.

PR 28-SEP-2000; 2000US-236032P.

PR 28-SEP-2000; 2000US-236033P.

PR 28-SEP-2000; 2000US-236034P.

PR 28-SEP-2000; 2000US-236109P.

PR 28-SEP-2000; 2000US-236111P.

PR 29-SEP-2000; 2000US-236842P.

PR 29-SEP-2000; 2000US-236891P.

PR 02-OCT-2000; 2000US-237122P.

PR 02-OCT-2000; 2000US-237173P.

PR 02-OCT-2000; 2000US-237278P.

PR 02-OCT-2000; 2000US-237294P.

PR 02-OCT-2000; 2000US-237295P.

PR 02-OCT-2000; 2000US-237316P.

PR 03-OCT-2000; 2000US-237425P.

PR 03-OCT-2000; 2000US-237598P.

PR 03-OCT-2000; 2000US-237604P.

PR 03-OCT-2000; 2000US-237606P.

PR 01-NOV-2000; 2000US-244867P.

PR 01-NOV-2000; 2000US-245084P.

XX

PA (AVAL-) AVALON PHARM.
 XX
 PI Young PE, Augustus M, Carter KC, Edner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX
 DR WPI; 2002-188264/24.

XX Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set
 XX
 PS Claim 1; SEQ ID 1009; 44pp; English.

XX The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL6164
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytosolic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilm's tumour.

XX Sequence 2691 BP; 679 A; 729 C; 695 G; 588 T; 0 other;

XX Query Match 3.0%; Score 47; DB 24; Length 2691;

XX Best Local Similarity 100.0%; Pred. 0.039; Mismatches 0; Indels 0; Gaps 0;

Oy 1507 GCTGGCCGTCGCTAGCTGCTGCGGCGTCGCCGCC 1553
 |||||||
 Db 1 GCTGGCCGTCGCTAGCTGCTGCGGCGTCGCCGCC 47

RESULT 15
 ABL199476/c
 ID ABL199476 standard; CDNA: 1907 BP.

AC ABL199476;

DT 07-MAR-2002 (first entry)

DE Mouse ischemic condition related CDNA sequence SEQ ID NO:433.

XX Mouse; ischemia; compressive ischaemia; occlusive ischaemia;
 KM vasospastic ischaemia; ischemic condition; ischaemic disease; ss.

XX Mus musculus.

OS WO200188188-A2.

PN 22-NOV-2001.

PR 18-MAY-2001; 2001WO-JP04192.

PR 18-MAY-2000; 2000JP-0145977.

PR (UYNT-) UNIV NIHON SCHOOL JURIDICAL PERSON.

XX Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;

XX WPI; 2002-034733/04.

DR P-Psdb; ABB57178.

XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring

PT expression levels of particular genes defined in the specification or
PT by determining the expression profile of a gene group comprising these
PT genes -

XX
PS Claim 2; Page 1188-1192; 2690pp; English.

CC The present invention describes a method for examining ischemic
CC conditions, comprising measuring the expression levels of particular
CC genes (1) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (1). The method
CC is useful for examining the ischemic condition (e.g. compressive
CC ischemia, occlusive ischemia or vasospastic ischemia) by measuring
CC expression levels of particular genes (AB199202 to AB199912, encoding
CC the protein sequences in AB199202 to AB199912) or by determining the
CC expression profile of a gene group comprising these genes. The
CC expression levels or expression profiles produced by these genes are
CC used as an indicator when screening for ischemic condition-improving
CC drugs or therapeutics for ischemic diseases. AB199913 and AB199914
CC represent PCR primers for a mouse ischemic condition related sequence,
CC which are used in the exemplification of the present invention.

XX
SQ Sequence 1907 BP; 379 A; 635 C; 600 G; 285 T; 8 other;

Query Match

Best Local Similarity 47.28; Score 46.6; DB 24; Length 1907;

Matches 128; Conservative 4; Mismatches 139; Indels 0; Gaps 0;

QY 1279 GGGCTGACAGTGAAGGCAAGGCTGGGAAATGAGACAGGCAAGTGGGAGCGGAT 1338
DB 399 GCGCTGTGCGCGGAGGCGAGCGCGGCGGAGGCGGTGGGAGTGGCGGAGGCGAC 340
QY 1339 GGGTCCCGAGGCGAGGCGGCGACAGGGTTGGAGAGCGCCAGCGCGCCCTTCCTT 1398
DB 339 GSCCGGGGAGGCGCGTGTGGAGGCGCGTGGGAGCGCGCGCCGCC 280
QY 1399 GCAAGGGCGCGCCAGCTTCCCGCCCTGCGCTCCCTCCCGCTGCGAGCTTACT 1458
DB 279 GCCGCGGACCTTCCCGCCCGCCCGCCCGCTTCCCGCCAGCCCGCGCTCC 220
QY 1459 TAACCTGACCGGCGGAGGCGGCTCTCACTTCCCTGAGCGCGCGCTTGCCTCG 1518
DB 219 GCGCGGTCTCTGCGCTGCAAGACGAGGCTGCGCTCCCGGTTCGCCCTTGCATCCC 160
QY 1519 GTCGCTAGCTCGCTCGGTGCGCGCTGCTCCG 1549
DB 159 AGSGCAGTCCGCGCTGCGGCGCGCGCG 129

Search completed: February 23, 2003, 06:41:03
Job time : 576 secs

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 23, 2003, 06:26:32 ; Search time 2083 Seconds
(without alignments)
12074.696 Million cell updates/sec

Title: US-09-865-879-2

Perfect score: 1553
Sequence: 1 gactcagggtctcccaacc.....ggtgcgctgctccgcctcc 1553

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	120	7.7	300	9	AU099192
2	92.4	5.9	793	9	AU119221
3	91.8	5.9	433	14	BM843883
4	91	5.9	620	10	BE336857
5	79	5.1	829	12	BG327480
6	79	5.1	842	12	BG775529

7	79	5.1	853	12	BG422226
8	79	5.1	895	12	AL546496
9	70.8	4.6	970	9	AL577405
10	70.2	4.5	987	17	CNS00418
11	67.2	4.3	1021	14	BQ434910
12	66.8	4.3	997	17	CNS005TE
13	66.6	4.3	910	17	CNS006ON
14	65.8	4.2	1101	17	CNS00397
15	65.2	4.2	966	14	BQ672441
16	65	4.2	799	9	AU120568
17	65	4.2	3172	11	BC026352
18	64.8	4.2	1016	14	BQ918616
19	64	4.1	975	12	BG425373
20	64	4.1	1315	13	BM543062
21	63.4	4.1	681	9	AU136140
22	63	4.1	1203	17	CNS015Y4
23	61.8	4.0	1050	14	BQ894806
24	61.6	4.0	1230	13	BM558820
25	60.8	3.9	814	17	AG062282
26	60.6	3.9	928	17	AG130881
27	60.6	3.9	985	14	BQ672571
28	60.6	3.9	1316	13	BM477175
29	60.4	3.9	932	17	CNS00720
30	59.8	3.9	961	17	AG083216
31	59.4	3.8	910	17	CNS006ON
32	59.4	3.8	1176	13	BM544862
33	59.2	3.8	918	12	BC809598
34	59.2	3.8	1364	14	BM810045
35	59	3.8	542	12	BG419864
36	59	3.8	885	12	BG423070
37	59	3.8	892	17	BH133618
38	59	3.8	924	12	BG420398
39	59	3.8	935	17	CNS006XK
40	59	3.8	982	12	BG419111
41	58.8	3.8	732	17	AZ196033
42	58.8	3.8	925	17	CNS0091P
43	58.8	3.8	980	14	BQ228348
44	58.4	3.8	867	12	BG423038
45	58	3.7	925	17	CNS0091P

ALIGNMENTS

RESULT 1
LOCUS AU099192 300 bp mRNA linear EST 05-APR-2001
DEFINITION AU099192 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone HRC01126 similar to Human transforming growth factor-beta induced gene product (BIG3) mRNA, mRNA sequence.

ACCESSION AU099192
VERSION AU099192.1 GI:13550321
KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 300)
Suzuki, Y., Tsunoda, T., Taira, H., Mizushima-Sugano, J., Sese, J., Hata, H., Oca, T., Isogai, T., Tanaka, T., Nakamura, Y., Mochizuki, S., Okubo, K., Suyama, A. and Sugano, S.

TITLE

In silico mapping of the 5'-ends of human mRNAs using full-length enriched and 5'-end enriched cDNA libraries constructed by

Oligo-capping method
Unpublished (2001)

JOURNAL COMMENT

Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp
Suzuki, Y., Yoshitomo-Nakagawa, K., Matuyama, K., Suyama, A. and Sugano, S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES	SOURCE	LOCATION/QUALIFIERS
		1. 300
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/clone="HRC01126"
		/clone_1fb="Sugano Homo sapiens cDNA library"
		/note="Differential display comparison of untreated and dimethylformate treated U937 cells"
BASE COUNT		32 a 117 c 95 g 56 t
ORIGIN		

Query Match	7.7%	Score	120	DB	9	Length	300
Best Local Similarity	100.0%	Pred. No.	3	7e-19			
Matches	120	Conservative	0	Mismatches	0	Indels	0
						Gaps	0

QY 1434 CGCTCCCTCCCGCTCCAGCTTACTTAACCTGGCCCGGGCGAGGGCGCTTCACTTTC 149

Db 1 CGCTCCCTCCCGCTCCAGCTTACTTAACCTGGCCCGGGCGAGGGCGCTTCACTTTC 60

QY 1494 CCTGGAGCCGCCCGCTTGGCCGTCGGTCGCTAGCTCGCTCGGTGCGGCGCTGTCGGCTCC 155
|||||
61 CCGGAGCCGCCCGCTTGGCCGTCGGTCGCTAGCTCGCTCGGTGCGGCGCTGTCGGCTCC 120
Db

RESULT 2	LOCUS	DEFINITION	793 bp	mrna	linear	EST 01-AUG-2002
AU119221	AU119221	HEMBA1 Homo sapiens cDNA clone HEMBA1005307 5', mRNA				

ACCESSION	AU119221	
VERSION	AU119221.1	GI:10934456
KEYWORDS	EST.	
SOURCE	human.	

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominae; Homo. 1 (bases 1 to 793)

Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,

AUTHORS

TITLE	HRI human cDNA project
JOURNAL	Unpublished (2000)
COMMENT	Contact: Takao Isogai

Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix Research Institute; cDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

FEATURES	Location/Qualifiers
source	1. .793

BASE COUNT	161	a	256	c	216	g	159	t	1	others
ORIGIN										
	/tissue_type="whole embryo, mainly head"									
	/dev_stage="embryo, 10 weeks"									
	/note="Vector: PME18FL3"									

Query Match	5.9%	Score	92.4	DB	9	Length	733
Best Local	98.9%	Pred. No.	3.8e-12				
Matches	93	Conservative	0	Mismatches	1	Indels	0
						Gaps	0

Dy 1460 AACCTGCCCCGGCGGGAGGCCCTCTCACTTCCCTGGAGCGCCGCCCTTGCCCGTGG 1519
Dd - 1 AACCTGCCCCGGCGGGAGGCCGTCCACACTTCCTGGAGCGCCGCCCTTGCCCGTGG 60

QY 1520 TCGTAGCTCGCTCGGTGCGCGCTGTCCCGCTCC 1553
|||||
Db 61 TCGTAGCTCGCTCGGTGCGCGCTGTCCCGCTCC 94

LOCUS	BM843883	433 bp	mRNA	linear	EST 06-MAR-2002
DEFINITION	K-ES0121878	S12SNU216	Homo sapiens	CDNA clone	S12SNU216-58-c05 5',
	mRNA sequence.				

ORGANISM	human.
SOURCE	EST.
KEYWORDS	BM843883.1
VERSION	GI:19200292
ACCESSION	BM843883

REFERENCE
1 (bases 1 to 433)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Oh, K. J., Cheong, J. E., Sohn, H. Y., Kim, J. M., Park, H. S., Kim, S. and Kim, Y. S.
21C Frontier Korean EST Project 2001

Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470

FEATURES

Email: Yongsung@email.krabb.re.kr
Plate: 58 row: G column: 05
High quality sequence stop: 433
Location/Qualifiers

```

/cell_type="Epithelial"
/cell_line="SNU-216"
/lab_host="Top10F"

```

I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with *E. coli* DNA ligase after digestion of

EcORI which site is converted to a DNA s

BASE COUNT	76 a	146 c	128 g	82 t	1 others
ORIGIN					
Query Match		5.58;	Score 91.8;	DB 14;	Length 433
Best Local Similarity		97.5%;	Pred. No. 4.3e-12;		

QY 1459 TAACTGGCCCCGCGGAGGACGCTTCACACTTCCCTGGAGGCCGCCGCTTGCCCCGCG 1518
| | |||||

QY 1519 GTCGCTAGCTTCGCTCGGTCGGCGTGTCCCGCTCC 1553
|||||

RESULT 4
BE336857
LOCUS
DEFINITION BE336857 620 bp mRNA linear EST 14-JUN-2000
ba97605.y1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:2957889 5
similar to gb:U19932 Mouse (MOUSE);, mRNA sequence.
ACCESSION BE336857
VERSION BE336857.1 GI:9189242
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mhc.nci.nih.gov/
1 (bases 1 to 620)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Other ESTs: ba97105.x1
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/resources.shtml
Seq primer: -40RP from Gibco
High quality sequence stop: 461.
Location/Qualifiers
1. 620
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2957889"
/clone_1lb="NIH_MGC_14"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
ECORI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 113 a 207 c 181 g 117 t 2 others
ORIGIN
Query Match 5.9%; Score 91; DB 10; Length 620;
Best Local Similarity 100.0%; Pred. No. 7.8e-12;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1463 CTGCGCCGGGGCGGAGGCGCTCTCACTTCCTGAGCCGCCGCTGCGCTCGTCG 1522
|||||
DB 1 CTGCGCCGGGGCGGAGGCGCTCTCACTTCCTGAGCGCCGCTGCGCTCGTCG 60
OY 1523 CTAGCTCGCTCGGTGGCGCTGCTCCGCTCC 1553
|||||
DB 61 CTAGCTCGCTCGGTGGCGCTGCTCCGCTCC 91

RESULT 5
BG327480
LOCUS
DEFINITION BG327480 829 bp mRNA linear EST 27-FEB-2001
602426462F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4564383 5',
mRNA sequence.
ACCESSION BG327480
VERSION BG327480.1 GI:13133918
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 829)
TITLE NIH-MGC http://mhc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM4279 row: 1 column: 16
High quality sequence stop: 704.
Location/Qualifiers
1. 829
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4564383"
/clone_1lb="NIH_MGC_14"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
ECORI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 172 a 268 c 233 g 156 t
ORIGIN
Query Match 5.1%; Score 79; DB 12; Length 829;
Best Local Similarity 100.0%; Pred. No. 8.3e-09;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1475 GCGGAGCGCTCTCACTTCCTGAGCGCCGCCGCTGCGCTGAGCTCGTCG 1534
|||||
DB 1 GCGGAGCGCTCTCACTTCCTGAGCGCCGCCGCTGCGCTGAGCTCGTCG 60
OY 1535 GTGCGCTCGTCCCGCTCC 1553
|||||
DB 61 GTGCGCTCGTCCCGCTCC 79

RESULT 6
BG775529
LOCUS
DEFINITION BG775529 842 bp mRNA linear EST 15-MAY-2001
602650440F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:4761465 5',
mRNA sequence.
ACCESSION BG775529
VERSION BG775529.1 GI:14045846
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mhc.nci.nih.gov/
1 (bases 1 to 842)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: L1CM1613 row: p column: 10
High quality sequence stop: 787.
Location/Qualifiers
1. .842

FEATURES

source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_40"
/issue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pOTB; Site: 1: XhoI;
Site: 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(g). Library constructed by
Liang Hong in the Laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

BASE COUNT 179 a 269 c 230 g 164 t

ORIGIN

Query Match 5.1%; Score 79; DB 12; Length 842;
Best Local Similarity 100.0%; Pred. No. 8.3e-09;

Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1475 GCGGAGCGCTCTACTTCCCTGGAGCCGCCGCTTCCCGTGGCTAGCTCGCTCG 1534

DB 1 GCGGAGCGCGCTCTACTTCCCTGGAGCCGCCGCTTCCCGTGGCTAGCTCGCTCG 60

QY 1535 GTGCGCGTCTCCGCTCC 1553

DB 61 GTGCGCGTCTCCGCTCC 79

RESULT 7 853 bp mRNA linear EST 14-MAR-2001
LOCUS BG422226
DEFINITION 602446857F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4585263 5',
mRNA sequence.

ACCESSION BG422226
VERSION BG422226.1 GI:13328732
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 853)
NIH-MGC http://mhc.ncl.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP

CDNA Library Preparation: Liang Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incey Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://lmgc.lnl.gov
Plate: L1CM1312 row: b column: 16
High quality sequence stop: 709.
Location/Qualifiers
1. .853

FEATURES

source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_14"
/issue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: pOTB; Site: 1: XhoI; Site: 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(g). Size selected >500bp for average

Insert size 1.8kb. Library constructed by Liang Hong in
the Laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 179 a 269 c 238 g 167 t

ORIGIN

Query Match 5.1%; Score 79; DB 12; Length 853;
Best Local Similarity 100.0%; Pred. No. 8.4e-09;

Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1475 GCGGAGCGCTCTACTTCCCTGGAGCCGCCGCTTCCCGTGGCTAGCTCGCTCG 1534

DB 1 GCGGAGCGCGCTCTACTTCCCTGGAGCCGCCGCTTCCCGTGGCTAGCTCGCTCG 60

QY 1535 GTGCGCGTCTCCGCTCC 1553

DB 61 GTGCGCGTCTCCGCTCC 79

RESULT 8 895 bp mRNA linear EST 16-FEB-2001
LOCUS AL546496
DEFINITION AL546496 LTI_NFL006_P12 Homo sapiens cDNA clone CS0D1030YC01 5
prime, mRNA sequence.

ACCESSION AL546496
VERSION AL546496.1 GI:12879668
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 895)
L.M.B., Gruber, C., Jesse, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.

REFERENCE 1
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source
1. .895
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="LTI_NFL006_P12"
/issue_type="placenta"
/note="Vector: pCMVSPORT 6; Site: 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact: Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifestech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 189 a 285 c 245 g 174 t 2 others

ORIGIN

Query Match 5.1%; Score 79; DB 9; Length 895;
Best Local Similarity 100.0%; Pred. No. 8.5e-09;

Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1475 GCGGAGCGCTCTACTTCCCTGGAGCCGCCGCTTCCCGTGGCTAGCTCGCTCG 1534

DB 1 GCGGAGCGCGCTCTACTTCCCTGGAGCCGCCGCTTCCCGTGGCTAGCTCGCTCG 60

QY 1535 GTGCGCGTCTCCGCTCC 1553

DB 61 GTGCGCGTCTCCGCTCC 79

RESULT 9
 AL577405 970 bp mRNA linear EST 16-FEB-2001
 LOCUS AL577405 LTI_NFL006.PL2 Homo sapiens cDNA clone CS001086YA21 5
 DEFINITION prime, mRNA sequence.
 ACCESSION AL577405
 VERSION AL577405.1 GI:12940501
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 970)
 AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 FEATURES
 source
 1..970
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS001086YA21"
 /clone_1lb="LTI_NFL006.PL2"
 /tissue_type="Placenta"
 /note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : filiang@lifetech.com URL : http://fulllength.invitrogen.com"
 BASE COUNT 208 a 303 c 258 g 194 t 7 others
 ORIGIN
 Query Match 4.6% Score 70.8; DB 9; Length 970;
 Best Local Similarity 97.3%; Pred. No. 9.5e-07;
 Matches 72; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Oy 1480 GCGCGCTACACTTCCCTGAGAGCGCCGCTTCCCGCTGATCGTACGTCGTCGTCGCG 1539
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 38 GGATGCTCAGCTTCCTGAGAGCGCCGCTTCCCGCTGATCGTACGTCGTCGTCGCG 97
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Oy 1540 CGTCGTCGCCGCTCC 1553
 ||||||||||||||||||
 Db 98 CGTCGTCGCCGCTCC 111
 ||||||||||||||||||
 RESULT 10
 CNS00418 987 bp DNA linear GSS 03-JUN-1999
 LOCUS CNS00418
 DEFINITION Drosophila melanogaster genome survey sequence TERT3 end of BAC #
 BACR03C16 of RPCI-96 library from Drosophila melanogaster (fruit
 fly) genomic survey sequence.
 ACCESSION AL066537
 VERSION AL066537.1 GI:4942778
 KEYWORDS GSS.
 SOURCE Drosophila melanogaster.
 ORGANISM Drosophila melanogaster.
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephyridroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 987)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)

[illegible]

Source

1.910

```

/db_xref="taxon:/22/
/clone="BACR14J21"
/clone_lib="RPCI-98"
/note="end : T7"

```

202. a	63 c	112 g	198 t	335 others .
--------	------	-------	-------	--------------

4.38;	Score 66.6;	DB 17;	Length 910;
Similarity 22.78;	Pred. No. 1e-05;		
Conservative 125;	Mismatches 117;	Indels 0;	Gaps 0

TTGACGACAGACTCTGGCGGAGGAGAGAGAGGAGGAGGCGTGCACGTACGCCAAG 1298
- - - - - :
CTGTCCTTTTTTTTGTCGCGGGGGGGKGGGGKGGGGGGCGHGGGGGGCCTD 498

[illegible][illegible]

CGGYSSGGBSSCCSTBTSSSTBTBKBTSTYSTSYTSTTTSTTSBBSYGC 678

[illegible]

```
|:::|
TBYTKST 751
```

00397 1101 bp DNA GSS 03-JUN-1999
 008A10 of *Phylla melanogaster* genome survey sequence T7 end of BAC #
 008A10 of RPCR-98 library from *Drosophila melanogaster* (fruit
 genomic survey sequence).

53912 GI:4941769

Phyla: Echinodermata, Cephalopoda, Mollusca, Annelida, Arthropoda, Hexapoda, Insecta; Pterygota: Ornithoptera, Neuroptera, Hymenoptera, Diptera, Braconyera, Muscomorpha, Coleoptera, Lepidoptera, Drosophilidae, Drosophila, bases 1 to 1101)

FEATURES

Location/Qualifiers

```

/orga
/db_x
/clor
/clor
/clor

```

111 a	286 c	221 q	241 t	242 others
-------	-------	-------	-------	------------

4.28; Score 65.8; DB 17; Length 1101;
Similarity 32.98; Pred. No. 1.7e-05;

GGCATGTTCTGGATCAGCACTGCCCTCCCCACCCCTCCCTGAGTGTG 1109
 :||::: || : ||| : :| :||: ||

CCTCAGATCAGTGAAGGAAATCTTCGGGCCCCAGCATGCAGAACC GAAGCCCCCA 1169
::: | :|: :: | ||: | | : | ||||:
ACCGGACATGTTTGAGGTGAGGTAAGAGGAGGAGGAGGAGGAGGAGGAGG AGC

AGCTGTCCCTCAGTCCCAGAAGGTCGCCCATTTTGGAAGCAGCCA CAGGAGGCCTAAGGA 1229
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
SCTCYCYCYGYGATCKYVTKRYKCRCTTYKBKRBTKGTSKRSGCGRSRGKGYYGG 896

ATTACCTTGTTGAGGAAGACTGTGCCAGGAGAGAGAGCGAAGGGCCTGGCACT 1289
| : : | : | ||| : : |||| | ||| | :
GGGGCCGYBSRGCGCTGKGC GCGCGCTCGWCGWSGGARGCGAGGGGGRRGGCSG R36

66CAAGGCGCTGGGAAAACGTAGCACCGGCACAGTGGGGAACGGGTGGTGCCAGG 1349
 ||| ||| ||| :||| ||| ||| ||| ||
 SGGGGSGGGGSGGGGCGGG 776

CCAGGGGCGCACGGGTGGAGGCCCAAGCGGGCCCCCTTCTTCACGGCCGG 1409
||| | ||| ||| | : : : : :
CGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGSGRASCAGAGCRSSGYC 716

1469
 1468
 1467
 1466
 1465
 1464
 1463
 1462
 1461
 1460
 1459
 1458
 1457
 1456
 1455
 1454
 1453
 1452
 1451
 1450
 1449
 1448
 1447
 1446
 1445
 1444
 1443
 1442
 1441
 1440
 1439
 1438
 1437
 1436
 1435
 1434
 1433
 1432
 1431
 1430
 1429
 1428
 1427
 1426
 1425
 1424
 1423
 1422
 1421
 1420
 1419
 1418
 1417
 1416
 1415
 1414
 1413
 1412
 1411
 1410
 1409
 1408
 1407
 1406
 1405
 1404
 1403
 1402
 1401
 1400
 1399
 1398
 1397
 1396
 1395
 1394
 1393
 1392
 1391
 1390
 1389
 1388
 1387
 1386
 1385
 1384
 1383
 1382
 1381
 1380
 1379
 1378
 1377
 1376
 1375
 1374
 1373
 1372
 1371
 1370
 1369
 1368
 1367
 1366
 1365
 1364
 1363
 1362
 1361
 1360
 1359
 1358
 1357
 1356
 1355
 1354
 1353
 1352
 1351
 1350
 1349
 1348
 1347
 1346
 1345
 1344
 1343
 1342
 1341
 1340
 1339
 1338
 1337
 1336
 1335
 1334
 1333
 1332
 1331
 1330
 1329
 1328
 1327
 1326
 1325
 1324
 1323
 1322
 1321
 1320
 1319
 1318
 1317
 1316
 1315
 1314
 1313
 1312
 1311
 1310
 1309
 1308
 1307
 1306
 1305
 1304
 1303
 1302
 1301
 1300
 1299
 1298
 1297
 1296
 1295
 1294
 1293
 1292
 1291
 1290
 1289
 1288
 1287
 1286
 1285
 1284
 1283
 1282
 1281
 1280
 1279
 1278
 1277
 1276
 1275
 1274
 1273
 1272
 1271
 1270
 1269
 1268
 1267
 1266
 1265
 1264
 1263
 1262
 1261
 1260
 1259
 1258
 1257
 1256
 1255
 1254
 1253
 1252
 1251
 1250
 1249
 1248
 1247
 1246
 1245
 1244
 1243
 1242
 1241
 1240
 1239
 1238
 1237
 1236
 1235
 1234
 1233
 1232
 1231
 1230
 1229
 1228
 1227
 1226
 1225
 1224
 1223
 1222
 1221
 1220
 1219
 1218
 1217
 1216
 1215
 1214
 1213
 1212
 1211
 1210
 1209
 1208
 1207
 1206
 1205
 1204
 1203
 1202
 1201
 1200
 1199
 1198
 1197
 1196
 1195
 1194
 1193
 1192
 1191
 1190
 1189
 1188
 1187
 1186
 1185
 1184
 1183
 1182
 1181
 1180
 1179
 1178
 1177
 1176
 1175
 1174
 1173
 1172
 1171
 1170
 1169
 1168
 1167
 1166
 1165
 1164
 1163
 1162
 1161
 1160
 1159
 1158
 1157
 1156
 1155
 1154
 1153
 1152
 1151
 1150
 1149
 1148
 1147
 1146
 1145
 1144
 1143
 1142
 1141
 1140
 1139
 1138
 1137
 1136
 1135
 1134
 1133
 1132
 1131
 1130
 1129
 1128
 1127
 1126
 1125
 1124
 1123
 1122
 1121
 1120
 1119
 1118
 1117
 1116
 1115
 1114
 1113
 1112
 1111
 1110
 1109
 1108
 1107
 1106
 1105
 1104
 1103
 1102
 1101
 1100
 1099
 1098
 1097
 1096
 1095
 1094
 1093
 1092
 1091
 1090
 1089
 1088
 1087
 1086
 1085
 1084
 1083
 1082
 1081
 1080
 1079
 1078
 1077
 1076
 1075
 1074
 1073
 1072
 1071
 1070
 1069
 1068
 1067
 1066
 1065
 1064
 1063
 1062
 1061
 1060
 1059
 1058
 1057
 1056
 1055
 1054
 1053
 1052
 1051
 1050
 1049
 1048
 1047
 1046
 1045
 1044
 1043
 1042
 1041
 1040
 1039
 1038
 1037
 1036
 1035
 1034
 1033
 1032
 1031
 1030
 1029
 1028
 1027
 1026
 1025
 1024
 1023
 1022
 1021
 1020
 1019
 1018
 1017
 1016
 1015

CG 651

6672441 966 bp mRNA linear EST 15-JUL-2001
 ENSECOURT-8349228 NIH_MGC_102 Homo sapiens CDNA clone IMAGE:6276244
 6672441 / mRNA sequence.
 6672441.1 GI:21783275

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 23, 2003, 06:33:27 ; Search time 87 Seconds
(without alignments)
5474.356 Million cell updates/sec

Title: US-09-865-879-2

Perfect score: 1553
Sequence: 1 gactcagggtgtcccaacc.....ggtgcgctgcgcgcctcc 1553

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : Issued Patents, NA: *
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/lna/PCRTUS.COMB.seq: *
6: /cgn2_6/ptodata/1/lna/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96.6	6.2	7218	1	US-08-232-463-14
2	47.6	3.1	7218	1	US-08-232-463-14
3	47.6	3.0	2691	1	US-07-878-960-1
4	46	3.0	53526	3	US-08-658-136-2
5	46	3.0	53577	3	US-08-658-136-1
6	45.4	2.9	1166	4	US-09-072-596-323
7	41.6	2.7	4897	6	5196516-7
8	41.4	2.7	2269	4	US-09-394-645-1
9	41.4	2.7	2269	4	US-09-243-560B-1
10	40.6	2.6	289	4	US-09-007-005-17
11	40.6	2.6	289	4	US-09-244-796-17
12	40	2.6	8438	1	US-07-945-283-1
13	39.6	2.5	3480	4	US-09-226-012-3
14	39.6	2.5	3950	4	US-09-226-012-1
15	39.2	2.5	320	4	US-09-165-264-7
16	39.2	2.5	2353	5	PCT-US92-06840-1
17	38.8	2.5	1719	4	US-09-330-740A-9
18	38.8	2.5	152331	3	US-09-128-155-16
19	38.6	2.5	351	4	US-09-220-528-7
20	38.6	2.5	351	4	US-09-220-528-10
21	38.6	2.5	423	4	US-09-220-528-8
22	38.6	2.5	423	4	US-09-220-528-11
23	38.6	2.5	546	4	US-09-220-528-42
24	38.6	2.5	663	4	US-09-220-528-24
25	38.6	2.5	663	4	US-09-220-528-25
26	38.6	2.5	663	4	US-09-220-528-44
27	38.6	2.5	696	4	US-09-220-528-1

C 28	38.6	2.5	696	4	US-09-220-528-2	Sequence 2, Appl
C 29	38.6	2.5	714	4	US-09-220-528-30	Sequence 30, Appl
C 30	38.6	2.5	714	4	US-09-220-528-31	Sequence 31, Appl
C 31	38.6	2.5	1652	4	US-09-220-528-68	Sequence 68, Appl
C 32	38.6	2.5	1652	4	US-09-220-528-69	Sequence 69, Appl
C 33	38.6	2.5	1763	6	5198542-1	Patent No. 5198542
C 34	38.4	2.5	1464	1	US-08-243-542-5	Sequence 5, Appl
C 35	38.4	2.5	1464	1	US-08-477-407-5	Sequence 5, Appl
C 36	38.4	2.5	1464	1	US-08-484-355-5	Sequence 5, Appl
C 37	38.4	2.5	2913	1	US-08-243-542-7	Sequence 7, Appl
C 38	38.4	2.5	2913	1	US-08-477-407-7	Sequence 7, Appl
C 39	38.4	2.5	2913	1	US-08-484-355-7	Sequence 7, Appl
C 40	38.4	2.5	2923	1	US-08-243-542-6	Sequence 6, Appl
C 41	38.4	2.5	2923	1	US-08-477-407-6	Sequence 6, Appl
C 42	38.4	2.5	2923	1	US-08-484-355-6	Sequence 6, Appl
C 43	38.4	2.5	3183	1	US-08-243-542-8	Sequence 8, Appl
C 44	38.4	2.5	3183	1	US-08-477-407-8	Sequence 8, Appl
C 45	38.4	2.5	3183	1	US-08-484-355-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMUG
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-F1s
US-08-232-463-14
Query Match 6.2%, Score 96.6; DB 1; Length 7218;

Best Local Similarity 7.08; Pred. No. 2.2e-16;
Matches 30; Conservative 254; Mismatches 143; Indels 0; Gaps 0;

```
OY 58 TCATTCGACATGGCAGCATGCTGCTGCAATGTCAGACCAAAACCTGACATAT 117
   | |||| | | |||| | | | | | | | | | | | | | | | | | | | | | |
Db 1028 TAATTCGACATGGCTGCGAGCTGAGGAGCTTGCATTTTTTTTTTTTTTTTT 1087
OY 118 CCTTGACTCTCTCTCTCCCTCTCCCTGTATACAGACTCCAAATCTATTTGAGACTAT 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1088 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1147
OY 178 TACCTCTACACCCCTCACATTTGCCAGCCTTCCCATCTCTCCCTTACACCATAGT 237
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1148 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1207
OY 238 TCAAGCTCCCAAGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 297
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1208 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1267
OY 298 TGACACTGCGCATGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 357
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1268 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1327
OY 358 GGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 417
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1328 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1387
OY 418 CAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 477
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1388 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1447
OY 478 AACTTTC 484
   | | | |
Db 1448 CTTCTC 1454
```

RESULT 2

US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300

TELEFAX: (703)683-4109
TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match 3.18; Score 47.6; DB 1; Length 7218;
Best Local Similarity 8.38; Pred. No. 0.0041;
Matches 17; Conservative 120; Mismatches 69; Indels 0; Gaps 0;

```
OY 1203 GGAAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1262
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1238 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1179
OY 1263 GCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1322
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1178 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1119
OY 1323 CAGTGGGAGCGGGTGGTGGCCAGGAGCAGGAGGAGGAGGAGGAGGAGGAG 1382
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1118 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1059
OY 1383 GCGGCGGCGCTCTGACAGGCGG 1408
   | | | | | | | | | |
Db 1058 CTCCTGACCTGACGAGGAGCTGG 1033
```

RESULT 3

US-07-878-960-1
; Sequence 1, Application US/07878960
; Patent No. 544164
; GENERAL INFORMATION:
; APPLICANT: Purchio, Anthony F.
; APPLICANT: Skonier, John
; APPLICANT: Neuhauer, Michael G.
; TITLE OF INVENTION: TGF-BETA INDUCED GENE AND PROTEIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bristol-Myers Squibb Company
; STREET: 3005 First Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/878,960
; FILING DATE: 05-MAY-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/833,835
; FILING DATE: 05-FEB-1992
; NAME: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Sorrentino, Joseph M.
; REGISTRATION NUMBER: 32,598
; REFERENCE/DOCKET NUMBER: ONO092-
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206/728-4800
; TELEFAX: 206/727-3601
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2691 base pairs

```

?      TYPE: nucleic acid
?      STRANDNESS: double
?      MOLECULE TYPE: linear
?      MOLECULE TYPE: CDNA
?      HYPOTHETICAL: NO
?      ORIGINAL SOURCE:
?      ORGANISM: Homo sapiens
?      TISSUE TYPE: LUNG
?      CELL TYPE: ADENOCARCINOMA
?      CELL LINE: A549
US-07-878-960-1

```

Query Match	3.0%;	Score 47;	DB 1;	Length 2691;
Best Local Similarity	100.0%;	Pred. No. 0.0039;		
Matches 47;	Conservative 0;	Mismatches 0;	Indels	

QY 1507 GCTTGCCCGTCGGTCGGCTAGCTCGCTCGGTGCGCGCTGTCGCCGCTCC 1553
|||||
Db 1 GCTTGCCCGTCGGTCGGCTAGCTCGCTCGGTGCGCGCTGTCGCCGCTCC 47

RESULT 4
US-08-658-136-2
; Sequence 2, Application US/08658136

1 GENERAL INFORMATION:
 2 APPLICANT: KLINGER, KATHERINE W
 3 APPLICANT: LANDES, GREGORY M
 4 APPLICANT: BURN, TIMOTHY C
 5 APPLICANT: CONNORS, TIMOTHY D
 6 APPLICANT: DACKOWSKI, WILLIAM
 7 APPLICANT: GERMINO, GREGORY
 8 APPLICANT: QIAN, FENG
 9 TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE
 10 NUMBER OF SEQUENCES: 58
 11 CORRESPONDENCE ADDRESS:
 12 ADDRESSEE: GENZYME CORPORATION
 13 STREET: ONE MOUNTAIN ROAD
 14 CITY: FRAMINGHAM
 15 STATE: MASSACHUSETTS
 16 COUNTRY: USA
 17 ZIP: 01701
 18 COMPUTER READABLE FORM:
 19 MEDIUM TYPE: Floppy disk
 20 COMPUTER: IBM PC compatible
 21 OPERATING SYSTEM: PC-DOS/MS-DOS
 22 SOFTWARE: Patent Release #1.0, Version #1.25
 23 CURRENT APPLICATION DATA:
 24 APPLICATION NUMBER: US/08/658,136

```

? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: LASEN, ELIZABETH
? REGISTRATION NUMBER: 31,845
? REFERENCE/DOCKET NUMBER: GEN4-17.8
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 508-872-8400
? TELEFAX: 508-872-5415
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 53526 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
IS-O8-658-136-2

```

	Query Match	3.0%;	Score 46;	DB 3;	Length 53526;
	Best Local Similarity	49.7%;	Pred. No. 0.026;		
	Matches 144;	Conservative	0;	Mismatches 145;	Indels 1;
0Y	126	CTCCTTCTCCCTCTTCCCTCTATACAGACTCCCAATTCATTATGAGACTATTACTCTCT	185		

[illegible]

RESULT 5
US-08-658-136-1
; Sequence 1, Application US/08658136

```

1      GENERAL INFORMATION:
2      APPLICANT: KLINGER, KATHERINE W
3      APPLICANT: LANDES, GREGORY M
4      APPLICANT: BURN, TIMOTHY C
5      APPLICANT: CONNORS, TIMOTHY D
6      APPLICANT: DACKOWSKI, WILLIAM
7      APPLICANT: DACKOWSKI, GREGORY
8      APPLICANT: QIAN, FENG
9      TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENB
10     NUMBER OF SEQUENCES: 58
11     CORRESPONDENCE ADDRESS:
12     ADDRESSEE: GENZYME CORPORATION
13     STREET: ONE MOUNTAIN ROAD
14     CITY: FRAMINGHAM
15     STATE: MASSACHUSETTS
16     COUNTRY: USA
17     ZIP: 01701
18     COMPUTER READABLE FORM:
19     MEDIUM TYPE: Floppy disk
20     COMPUTER: IBM PC compatible
21     OPERATING SYSTEM: PC-DOS/MS-DOS
22     SOFTWARE: PatentIn Release #1.0, Version #1.25
23     CURRENT APPLICATION DATA:
24     APPLICATION NUMBER: US/08/658,136

```

```

: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: LASSEN, ELIZABETH
: REGISTRATION NUMBER: 31,845
: REFERENCE/DOCKET NUMBER: GEN4-17, 8
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 508-872-8400
: TELEFAX: 508-872-5415
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 53577 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: OS-08-658-136-1

```

Query Match	3.0%;	Score 46;	DB 3;	Length 53577;
Best Local Similarity	49.7%;	Pred. NO. 0.026;		
Matches 144;	Conservative	0;	Mismatches 145;	Indels 1;
Gaps				
126	CTCCTTTCCTCCCTTCCTCCGTATACAGACGCCAAATTCATTTGAGACTATTAACCTCT	185		
	111 111 111 111 111 111 111 111 111 111 111 111 111 111 111			
DB 34583	CTCCCT	34642		

[illegible]

RESULT 6
UE-09-07

```

1 Sequence 323, Application US/09072596
2 Patent No. 6458366
3
4 GENERAL INFORMATION:
5 APPLICANT: Reed, Steven G.
6 APPLICANT: Skelky, Yasir A.W.
7 APPLICANT: Dillon, Davin C.
8 APPLICANT: Campos-Neto, Antonia
9 APPLICANT: Houghton, Raymond
10 APPLICANT: Vedvick, Thomas S.
11 APPLICANT: Twardzik, Daniel R.
12 APPLICANT: Lodes, Michael J.
13 APPLICANT: Hendrickson, Ronald C.
14 TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
15 TUBERCULOSIS
16 NUMBER OF SEQUENCES: 350
17 CORRESPONDENCE ADDRESS:
18 ADDRESSEE: SEED and BERRY LLP
19 STREET: 6300 Columbia Center, 701 Fifth Avenue
20 CITY: Seattle
21 STATE: Washington
22 COUNTRY: USA
23 ZIP: 98104-7092
24
25 COMPUTER READABLE FORM:
26 MEDIUM TYPE: Floppy disk
27 COMPUTER: IBM PC compatible
28 OPERATING SYSTEM: PC-DOS/MS-DOS
29 SOFTWARE: PatentIn Release #1.0, Version #1.30
30 CURRENT APPLICATION DATA:
31 APPLICATION NUMBER: US/09/072,596
32 FILING DATE: 05-MAY-1998
33 CLASSIFICATION:
34 ATTORNEY/AGENT INFORMATION:
35 NAME: Makl, David J.
36 REGISTRATION NUMBER: 31,392
37 REFERENCE/DOCKET NUMBER: 210121.417C9
38 TELECOMMUNICATION INFORMATION:
39 TELEPHONE: (206) 622-4900
40 TELEFAX: (206) 682-6031
41 INFORMATION FOR SEQ ID NO: 323:
42 SEQUENCE CHARACTERISTICS:
43 LENGTH: 1166 base pairs
44 TYPE: nucleic acid
45 STRANDEDNESS: single
46 TOPOLOGY: Linear
47
48 MOLECULE TYPE: Genomic DNA
49
50 US-09-072-596-323

```

Query Match	2.98;	Score 45.4;	DB 4;	Length 1166;
Best Local Similarity	28.98;	Pred. No. 0.0075;		
Matches	86;	Conservative 61;	Mismatches 151;	Indels 0;
			Gaps	0;

Oy 156 CTCGAATTTCTATTGAGACTATTACCTTCTACACCCCTACATTTGGCCAGCCTTCCCA 215
 ||||: : : : : | : ||| : : | : : : ||
 Db 467 CTCCKMTTCAYCTCMTCCGGTCCCTTMMNCCSCNCRYCTCANCNCTKSGKACACNA 526

Qy	216	TCGTGGCTCTACACAAAGTTCAAGTCTCCCATGGTCCCTCTGGTAGCCGTCT	275
Db	527	TMTCSACRCHCTTMCYKMSCKANKTTTCCCTCCMCCCTYTNMCCAMCMSCSTCTTTCMAAC	568
Qy	276	TCCTTGCCTCCCTTAAGCCTCTCTATGACACATGGCCATGTCACTATGGCCCTTCACCACTATACCCG	335
Db	587	TKRCCGGCKCKCMCHCTCTCCCAAYNMAACCKRTCYCMCMNYCYWCKCKRAGYKNNMC	646
Qy	336	CTTAGGCTTACGTGGAGTCGTGGGGCCCTGCTACCTTCCTCCCTCTTCCCTACCCCTTGA	395
Db	647	TCCMACGCTMYNTTCTCTCMCCCMKACCKNTTCTCWSCCCCCAKACMYCAMCMT	706
Qy	396	CTCCACCTCCCTGTGCTTCACGCAACAGATATCTTAGTTCGTGAATGATGCATGCC	453
Db	707	MTCCMCTCKRAGSGCCCTYCNMNTCCNMCMCHCTCMCTNNATCAACMNTTCTTCTCTCMAMT	764

RESULT 7

Patent No. 5196516-7
 APPLICANTS: SCHREUDS, CHRISTA S.; MEPPENLEITER, THOMAS C.
 SIMON, ARTUR J.; LUDAS, NOEMI; RZHA, HANNS J.
 TITLE OF INVENTION: PSEUDORABIES VIRUS VACCINE
 NUMBER OF SEQUENCES: 8
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/383, 833
 FILING DATE: 21-JUL-1989
 SEQ ID NO.: 7
 LENGTH: 4897
 5196516-7

Query Match	2.7%	Score 41.6;	DB 6;	Length 4897;
Best Local Similarity	53.0%;	Pred. No. 0.15;		
Matches	89;	Conservative	0;	Mismatches 79;
			Indels	0;
			Gaps	0;

QY 1216 GGAAGCCCTAAGGACCTATACCTCTGGTTTGAAGAAAGACTGTGGCGAGGGAGACAGGAG 1215
 |||||
 Db 4408 GGAAGGGGAGAGCGGACGGAAGGGGAGAGGAGCAGCGGGGAGAGGGAGACGAGCGGAGCGGAC 4349
 |||||
 QY 1216 GGAAGGCGCTGGCAGTGAAGAGCGCAAGGGCTGGGAAACTGACACAGGGGCACAGTCTCGGGACG 1335
 |||||
 Db 4348 GGGGGGGGAGAGGAGACGGAGGGGAGAGGAGACGAGAGGGGAGAGGCGCGGGGGGGGAGAG 4289
 |||||
 QY 1336 GGTGGTGGCCAGGAGCAGCCAGGGGGCGACAGGGTTGGAGGGCCAGG 1383
 |||||
 Db 4288 GGAACGAGGGGAGAGGACCGGGGGGGAGAGGAGACGAGGGGGAGAGG 4211
 |||||

RESULT 8

```

Sequence 1 Application US/09394645
Patent No. 6380371
GENERAL INFORMATION:
APPLICANT: Sassetti, Christopher M.
APPLICANT: Rosen, Steven R.
TITLE OF INVENTION: ENDOGLYCAN: A NOVEL PROTEIN HAVING SELECTIVE
TITLE OF INVENTION: LIGAND AND CHEMOKINE PRESENTATION ACTIVITY
FILE REFERENCE: 6510-122051
CURRENT APPLICATION NUMBER: US/09/394,645
CURRENT FILING DATE: 1999-09-13
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 1
LENGTH: 2269
TYPE: DNA
ORGANISM: human
US-09-394-645-1

```

Query Match 2.78; Score 41.4; DB 4; Length 2269;

Matches 87; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

OY 1310 CTGAGCAGGGGCACAGTGGCGGAGCGGGCTGGGTGCCCGCCAGGGCAGCCAGGGGGCGCACGGGT 1369


```

Db      1954 CTCGGCGGCGGGGTCTTCCCTTCGCCGCAGTTCAACAGGCGGGCTTGACCAACTCCC 2013
Oy      1370 TGGGAGGGGCCAGGCGGGCCCGCCTTCCTTGACAGGGGCGGCCAGCTTCCCCGCCCTGG 1429
Db      2014 TCACATCCCGCCCGAAGGGGCGAGGCTCAAAAGCCCGCTTGGCCCCGCTTCGCCCGCTGA 2073
Oy      1430 GGTCCGCTCCCTCCGCGTAGCTTACTTAACCTGGGCCGGG 1472
Db      2074 ACCCGGCCCCCGGGGCGGGCGCGGCTTCTGCGCCCGGG 2116

RESULT 9
US-09-243-560B-1
; Sequence 1, Application US/09243560B
; Patent No. 6395882
; GENERAL INFORMATION:
; APPLICANT: Rosen, Steven D.
; APPLICANT: Sassetti, Christopher M.
; TITLE OF INVENTION: No. 6395882el Selectin Ligands
; FILE REFERENCE: UCAL097U51
; CURRENT APPLICATION NUMBER: US/09/243, 560B
; PRIOR FILING DATE: 1999-02-03
; PRIOR APPLICATION NUMBER: US 60/074,389
; PRIOR FILING DATE: 1998-02-11
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2269
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-243-560B-1

Query Match          2.7%; Score 41.4; DB 4; Length 2269;
Best Local Similarity 53.4%; Pred.No. 0.12;
Matches 87; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

Oy      1310 CTGACACAGGGGCACAGTGTCCGGAGCGGGTGGGTGCCCCAGGCGCACGCGGT 1369
Db      1954 CTCGGCGGCGGGGTCTTCCCTTCGCCGCAGTTCAACAGGCGGGCTTGACCAACTCCC 2013
Oy      1370 TGGGAGGGGCCAGGCGGGCCCGCCTTCCTTGACAGGGGCGGCCAGCTTCCCCGCCCTGG 1429
Db      2014 TCACATCCCGCCCGAAGGGGCGAGGCTCAAAAGCCCGCTTGGCCCCGCTTCGCCCGCTGA 2073
Oy      1430 GTCCGCTCCCTCCGCGTAGCTTACTTAACCTGGGCCGGG 1472
Db      2074 ACCCGGCCCCCGGGGCGGGCGCGGCTTCTGCGCCCGGG 2116

RESULT 10
US-09-007-005-17/c
; Sequence 17, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihue
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; TITLE OF INVENTION: FUSIONS
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007,005B
; CURRENT FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence

```

[illegible]

Oy	166	TATTTAGCATTTTACCTCTTACACCCCTACATTTGGCCACACCTTCCCATCTGTGCTC	225
Db	145	YNYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYS	86
Oy	226	TACACCATTAAGTTCAGCTCTCCATGGTCCCTTCCTGGTATTCGTTC	274
Db	85	YNYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYS	37
RESULT 12			
US-07-945-283-1			
Sequence 1, Application US/07945283			
Patent No. 5352596			
GENERAL INFORMATION:			
APPLICANT: Cheung, Andrew K.			
APPLICANT: Wesley, Ronald D.			
TITLE OF INVENTION: Pseudorabies Virus Deletion Mutants			
TITLE OF INVENTION: Involving the EP0 and LTR Genes			
NUMBER OF SEQUENCES: 7			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Curtiss P. Ribando			
STREET: 1815 No. 5352596th University Street			
CITY: Peoria			
STATE: IL			
COUNTRY: USA			
ZIP: 61604			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: Patentln Release #1.0, Version #1.25			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/07/945,283			
FILING DATE: 19920911			
CLASSIFICATION: 424			
ATTORNEY/AGENT INFORMATION:			
NAME: Ribando, Curtiss P.			
REGISTRATION NUMBER: 27976			
TELECOMMUNICATION INFORMATION:			
TELEPHONE: 309-685-4011 ext. 513			
TELEFAX: 309-685-4128			
INFORMATION FOR SEQ ID NO: 1:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 8438 base pairs			
TYPE: NUCLEIC ACID			
STRANDEDNESS: double			
TOPOLOGY: linear			
MOLECULE TYPE: DNA (genomic)			
HYPOTHETICAL: NO			
ANTI-SENSE: NO			
ORIGINAL SOURCE:			
ORGANISM: Pseudorabies virus			
FEATURE:			
NAME/KEY: CDS			
LOCATION: 622..6495			
FEATURE:			
NAME/KEY: variation			
LOCATION: replace(1099, "g")			
FEATURE:			
NAME/KEY: variation			
LOCATION: replace(1267, "c")			
FEATURE:			
NAME/KEY: variation			
LOCATION: replace(1381, "c")			
FEATURE:			
NAME/KEY: variation			
LOCATION: replace(1566, "c")			
FEATURE:			
NAME/KEY: variation			
LOCATION: replace(7010, "g")			
US-07-945-283-1			

[illegible]

```

RESULT 13
US-09-226-012-1/c
Sequence 1, Application US/09226012
Patent No. 6207383
GENERAL INFORMATION:
APPLICANT: Keating, Mark T.
APPLICANT: Splawski, Igor
TITLE OF INVENTION: MUTATIONS IN AND GENOMIC STRUCTURE OF HERG - A LONG QT
FILE REFERENCE: 2323-136
CURRENT APPLICATION NUMBER: US/09/226,012
CURRENT FILING DATE: 1999-01-06
EARLIER APPLICATION NUMBER: 09/122,847
EARLIER FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 116
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 3480
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(3477)
US-09-226-012-1

```

Query Match	2.5%	Score 39.6	DB 4	Length 3480
Best Local Similarity	51.7%	Pred. No. 0.44	Mismatches 84	Indels 0
Matches	90	Conservative	0	Gaps 0
QY 1280	GCGTGCAGTGAAGGGCGCTGGGAAAAC	TGAACACAGGCGACACATCTGCGGAGCGG	GTG	1339
Db 2900	GCGTCTCCGGGGGCGCTGGGGCTGGAGAA	GGGCAACACAGCGGAGGGGGCTGAGACTGCGG		2841
QY 1340	GGTCCACAGGGCACAGGGGGCGCACAGG	GTGAGAGCGCCAGCGGCCCGCCCTCTCTTG		1399
Db 2840	CTTGGGCGCTCACTCACTGCTCTCAGGGG	CGTGGAGGGGCGACACTGAGCGGGGCTCTCCCGC		2781
QY 1400	CAGGGGCGGGCCACAGTCCCGCGGCCCG	GGGGGTCCCGCTCCCTCCCGCTGCACAGC		1453
Db 2780	CAGGGCGCCCGCGGGCGCGCCCGGCTACT	CGGGCCCTTGCCCCCGCGGCCCGCGC		2727

```

RESULT 14
US-09-226-012-3/C
: Sequence 3, Application US/09226012
: Patent No. 6207383
: GENERAL INFORMATION:
: APPLICANT: Keating, Mark T.
: APPLICANT: Splawski, Igor
: TITLE OF INVENTION: MUTATIONS IN AND GENOMIC STRUCTURE OF HERG - A LONG QT
: TITLE OF INVENTION: SYNDROME GENE
: FILE REFERENCE: 2323-136
: CURRENT APPLICATION NUMBER: US/09/226,012
: CURRENT FILING DATE: 1999-01-06
: EARLIER APPLICATION NUMBER: 09/122,847
: EARLIER FILING DATE: 1998-07-27

```

```
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 3
; LENGTH: 3950
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (67)..(3543)
US-09-226-012-3
```

```
Query Match          2.5%; Score 39.6; DB 4; Length 3950;
Best Local Similarity 51.7%; Pred.No.0.46;
Matches 90; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
```

```
QY 1280 GGCCTGCAGTGAAGGCGCTGGAAAACTGACACGCGGACAGTGCAGGCGGTG 1339
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2966 GGCCTGCAGGCGGCGCTGGAGAAAGGACGAGGCGGAGGCGTGGAGCTGCGG 2907

QY 1340 GGTGCCAGGGGCAAGCGAGGGCGACGGGTTGGAGAGCGCCGCCCTCTTG 1399
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2906 CCTGGGCCCTCATCTCAGGCTGAGGCGGCACTGAGACGGGCTCTCCCC 2847

QY 1400 CAGGGCGGCGCCAGCTTCCCGCCCTGCGCTCCCTCCGCTGCGACG 1453
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2846 CAGGGCGGCGGCGGCGGCGGCTAGCGCTGCGCCCGCGCGCGCGGCG 2793
```

RESULT 15

```
US-09-165-264-7
; Sequence 7, Application US/09165264
; Patent No. 6197510
; GENERAL INFORMATION:
; APPLICANT: Vinayagamorthy, Thuraiayah
; TITLE OF INVENTION: Multi-loci Genomic Analysis
; FILE REFERENCE: 44747
; CURRENT APPLICATION NUMBER: US/09/165,264
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 7
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Primer sequence
US-09-165-264-7
```

```
Query Match          2.5%; Score 39.2; DB 4; Length 320;
Best Local Similarity 53.2%; Pred.No.0.21;
Matches 83; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
```

```
QY 1241 GGTTCAGGAAGCTGTGGCAGAGAGAGAGGAGGCGCTGAGTGAAGGCAAGG 1300
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 158 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 217

QY 1301 CTTCGAAAACTGACAGCGGCGACAGTGCAGGAGCGGCGGTGCTCCAGCGCAGGCG 1360
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 218 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 277

QY 1361 CGCAGCGGTTGGAGGCGCCAGGCGCCGCCCTCC 1396
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 278 GGGGGGGGGGGGGGGGGGGGGGGAGGCTAGCCATGC 313
```

Search completed: February 23, 2003, 08:25:08
Job time : 226 secs

GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 23, 2003, 06:41:08 ; Search time 142 Seconds
(without alignments)
6142.437 Million cell updates/sec

Title: US-09-865-879-2

Perfect score: 1553
Sequence: 1 gactcagggtgtcccaaac.....ggtagcgctgcgcgcctcc 1553

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 44218 segs, 280819700 residues

Total number of hits satisfying chosen parameters: 884236

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published_Applications_NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCOT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCOT_NEW_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	3.3	2782	12	US-10-044-090-647
2	47.6	3.1	407	9	US-10-025-380-175
3	47.6	3.1	407	10	US-09-922-217-175
4	47.6	3.1	407	10	US-09-833-263-175
5	47	3.0	2691	9	US-10-025-380-121
6	47	3.0	2691	10	US-09-922-217-121
7	47	3.0	2691	10	US-09-833-263-121
8	47	3.0	2691	10	US-09-880-107-2399
9	46.2	3.0	174424	10	US-09-967-768A-114
10	46	3.0	53522	9	US-09-904-968A-1
11	42.8	2.8	3441	9	US-09-944-413-6
12	42.8	2.8	3441	9	US-09-944-403-6
13	42.8	2.8	3441	9	US-09-944-896-6
14	42.8	2.8	3441	9	US-09-944-944-6
15	42.8	2.8	3441	9	US-09-944-907-6
16	42.8	2.8	3441	9	US-09-944-929-6
17	42.8	2.8	3441	10	US-09-866-028-6
18	42.8	2.8	3441	10	US-09-944-449-6
19	42.8	2.8	3441	10	US-09-944-457-6

20	42.8	2.8	3441	10	US-09-945-587-6	Sequence 6, Appl1
21	42.8	2.8	3441	10	US-09-945-015-6	Sequence 6, Appl1
22	42.8	2.8	3441	10	US-09-944-396-6	Sequence 6, Appl1
23	42.8	2.8	3441	10	US-09-944-097-6	Sequence 6, Appl1
24	42.8	2.8	3441	10	US-09-944-432-6	Sequence 6, Appl1
25	42.8	2.8	3441	10	US-09-943-762-6	Sequence 6, Appl1
26	42.8	2.8	3441	10	US-09-944-654-6	Sequence 6, Appl1
27	42.8	2.8	3441	10	US-09-943-851A-6	Sequence 6, Appl1
28	42.4	2.7	5025	10	US-09-960-251-16	Sequence 176, App
29	42.4	2.7	42999	9	US-09-799-462A-17	Sequence 17, Appl
30	42.4	2.7	42999	9	US-10-125-767-17	Sequence 17, Appl
31	42.4	2.7	42999	9	US-09-836-911A-17	Sequence 17, Appl
32	41.8	2.7	312	10	US-09-864-761-18121	Sequence 18121, A
33	41.8	2.7	1668	10	US-09-948-018-1	Sequence 3, Appl1
34	41.8	2.7	2223	9	US-09-999-248-3	Sequence 1, Appl1
35	41.8	2.7	88421	9	US-09-976-059-1	Sequence 1, Appl1
36	41.4	2.7	2269	10	US-09-860-298-1	Sequence 1, Appl1
37	41.4	2.7	8280	10	US-09-764-878-384	Sequence 384, App
38	41.4	2.7	1835	10	US-09-764-878-385	Sequence 385, App
39	41.2	2.7	492	9	US-09-854-133-323	Sequence 323, App
40	41.2	2.7	1079	10	US-09-738-973-323	Sequence 34, Appl
41	41.2	2.7	492	10	US-09-804-682-34	Sequence 3, Appl1
42	40.8	2.6	2351	10	US-09-803-589-3	Sequence 724, App
43	40.8	2.6	2500	10	US-09-954-456-724	Sequence 1169, Ap
44	40.8	2.6	2500	10	US-09-954-456-1169	Sequence 1827, Ap
45	40.8	2.6	2500	10	US-09-954-456-1827	

ALIGNMENTS

RESULT 1
US-10-044-090-647
; Sequence 647, Application US/10044090
; Patient No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Handman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044, 090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 647
; LENGTH: 2782
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 1519595CB1
US-10-044-090-647

Query Match 3.3%; Score 52; DB 12; Length 2782;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1502 CGCCCGCTTCCTCCGCTCGCTAGCTAGCTGCTGCGCGTCCGCTCC 1553
DB 1 CGCCCGCTTCCTCCGCTCGCTAGCTAGCTGCTGCGCGTCCGCTCC 52

RESULT 2
US-10-025-380-175
; Sequence 175, Application US/10025380
; Publication No. US20020182191A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong

DB 39270 CCCCCGGGACCCCGGAGCTCAGCGGAGCCAGCCAGCGGCTCCGCGAGACC 39329
QY 1540 CGTGTCTCCGC 1550
DB 39330 CCCTTCCCGC 39340

RESULT 10
US-09-904-968A-1
; Sequence 1, Application US/09904968A
; Publication No. US20030008288A1
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: GERMINO, Gregory
; APPLICANT: WATNICK, Terry
; APPLICANT: PHAKDEKITCHAROEN, Bunyong
; TITLE OF INVENTION: DETECTION AND TREATMENT OF POLYCYSTIC KIDNEY DISEASE
; FILE REFERENCE: JHU1680-2
; CURRENT APPLICATION NUMBER: US/09/904,968A
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/283,691
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/218,261
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 53522
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-904-968A-1

Query Match 3.0%; Score 46; DB 9; Length 53522;
Best Local Similarity 49.7%; Pred. No. 0.054;
Matches 144; Conservative 0; Mismatches 145; Indels 1; Gaps 1;

QY 126 CTCCTTTCCTCCCTTCCTCCCTGATACAGACTCCAAATTTATGAGACTATTACTCCT 185
DB 34584 CTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 34643
QY 186 ACACCCCTACATTTGCGCCAGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 245
DB 34644 TCCCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 34703
QY 246 TCCCATGATGCTCCCTCCCTGATACCTGTTCTTCTGCTCCCTTAAGCCCTCATGACACTG 305
DB 34704 CCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 34763
QY 306 GGCATGTCACCTGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTG 364
DB 34764 CCCATCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 34823
QY 365 CTACCTTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 414
DB 34824 CTGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 34873

RESULT 11
US-09-944-413-6
; Sequence 6, Application US/09944413
; Patent No. US20020156004A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Botstein, David
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Geritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul
; APPLICANT: Grimaldi, Christopher
; APPLICANT: Gurney, Austin

APPLICANT: Hillan, Kenneth
APPLICANT: Kijavijn, Ivar
APPLICANT: Napier, Mary
APPLICANT: Roy, Margaret
APPLICANT: Tumas, Daniel
APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
ACIDS ENCODING THE SAME
FILE REFERENCE: P2548P1C1
CURRENT APPLICATION NUMBER: US/09/944,413
CURRENT FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 09/866,028
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 60/067,411
PRIOR FILING DATE: December 3, 1997
PRIOR APPLICATION NUMBER: 60/069,334
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,335
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,278
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,425
PRIOR FILING DATE: December 12, 1997
PRIOR APPLICATION NUMBER: 60/069,696
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,694
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,702
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,870
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/069,873
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/068,017
PRIOR FILING DATE: December 18, 1997
PRIOR APPLICATION NUMBER: 60/070,440
PRIOR FILING DATE: January 5, 1998
PRIOR APPLICATION NUMBER: 60/074,086
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/074,092
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/075,945
PRIOR FILING DATE: February 25, 1998
PRIOR APPLICATION NUMBER: 60/112,850
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 60/113,296
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 60/146,222
PRIOR FILING DATE: July 28, 1999
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: September 16, 1998
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: December 1, 1998
PRIOR APPLICATION NUMBER: 09/216,021
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 09/218,517
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 09/254,311
PRIOR FILING DATE: March 3, 1999
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: June 22, 1999
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: September 15, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28409
PRIOR FILING DATE: NO. US20020156004A1
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: NO. US20020156004A1
PRIOR APPLICATION NUMBER: PCT/US99/28301
PRIOR FILING DATE: December 1, 1999
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: December 16, 1999
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: February 11, 2000


```

? PRIOR APPLICATION NUMBER: PCT/US00/04414
? PRIOR FILING DATE: February 22, 2000
? PRIOR APPLICATION NUMBER: PCT/US00/05841
? PRIOR FILING DATE: March 2, 2000
? PRIOR APPLICATION NUMBER: PCT/US00/08439
? PRIOR FILING DATE: March 30, 2000
? PRIOR APPLICATION NUMBER: PCT/US00/14042
? PRIOR FILING DATE: May 22, 2000
? PRIOR APPLICATION NUMBER: PCT/US00/20710
? PRIOR FILING DATE: July 28, 2000
? PRIOR APPLICATION NUMBER: PCT/US00/32678
? PRIOR FILING DATE: December 1, 2000
? PRIOR APPLICATION NUMBER: PCT/US01/06520
? PRIOR FILING DATE: February 28, 2001
? NUMBER OF SEQ ID NOS: 120
? SEQ ID NO 6
? LENGTH: 3441
? TYPE: DNA
? ORGANISM: Homo Sapien
US-09-944-413-6

```

LENGTH: 3441
TYPE: DNA
ORGANISM: Homo Sapien
US-09-944-403-6

Query Match 2.8%; Score 42.8; DB 9; Length 3441;
Best Local Similarity 51.0%; Pred. No. 0.096;
Matches 101; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 1325 GTGCGGAGCGGCGTGTGCGCCAGGCGCAGCGGCGTGGAGAGCGCCAGGC 1384
DB 8 GTGGGCGGAGCGGTGGGCGCCGCGCCAGCGCCCGCCGCGCCCTCCGCTCCGCACTC 67
QY 1385 GGGCGCGCCCTCTTGCAGGAGCGCGCCAGCTCCCGCCCTGCGCTCCCTCC 1444
DB 68 GCGCTCTCCCTCCCTCCCGCCGCTCCCGCTCCCTCCCTCCCTCCCTCCAGCTGTC 127
QY 1445 GCTCGAGCTTACTTAACCTGCGCCGCGCGCGCGCGCTCTCACTTCCCTGAGCGCC 1504
DB 128 GTTCGGGTATGCGAGAGCTCCGCGCCCGCGCGCGCGCTGCTCTCTGCGCTGCTG 187
QY 1505 CCGCTTGCGCCGCTGCGTGC 1522
DB 188 CTGCTGCGCTCCGCGCG 205

RESULT 13
Sequence 6, Application US/09444896
Patent No. US20020168715A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Botstein, David
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerlitsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul
APPLICANT: Grimaldi, Christopher
APPLICANT: Gurney, Austin
APPLICANT: Hillan, Kenneth
APPLICANT: Kijavlin, Ivar
APPLICANT: Napier, Mary
APPLICANT: Roy, Margaret
APPLICANT: Tumas, Daniel
APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P2548P1C1
CURRENT APPLICATION NUMBER: US/09/944, 896
CURRENT FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: 09/866, 028
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 60/069, 334
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069, 335
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069, 278
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069, 425
PRIOR FILING DATE: December 12, 1997
PRIOR APPLICATION NUMBER: 60/069, 696
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069, 694
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069, 702
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069, 870
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/069, 873
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/068, 017

PRIOR FILING DATE: December 18, 1997
PRIOR APPLICATION NUMBER: 60/070, 440
PRIOR FILING DATE: January 5, 1998
PRIOR APPLICATION NUMBER: 60/074, 086
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/074, 092
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/075, 945
PRIOR FILING DATE: February 25, 1998
PRIOR APPLICATION NUMBER: 60/112, 850
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 60/113, 286
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 60/146, 222
PRIOR FILING DATE: July 28, 1999
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: September 16, 1998
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: December 1, 1998
PRIOR APPLICATION NUMBER: 09/216, 021
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 09/218, 517
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 09/254, 311
PRIOR FILING DATE: March 3, 1999
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: June 22, 1999
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: September 15, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28409
PRIOR FILING DATE: No. US20020168715A1ember 30, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: No. US20020168715A1ember 30, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28301
PRIOR FILING DATE: December 1, 1999
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: December 16, 1999
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: February 11, 2000
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: February 22, 2000
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: March 2, 2000
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: March 30, 2000
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: May 22, 2000
PRIOR APPLICATION NUMBER: PCT/US00/20710
PRIOR FILING DATE: July 28, 2000
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: December 1, 2000
PRIOR APPLICATION NUMBER: PCT/US01/06520
PRIOR FILING DATE: February 28, 2001
NUMBER OF SEQ ID NOS: 120
SEQ ID NO 6
LENGTH: 3441
TYPE: DNA
ORGANISM: Homo Sapien
US-09-944-896-6
Query Match 2.8%; Score 42.8; DB 9; Length 3441;
Best Local Similarity 51.0%; Pred. No. 0.096;
Matches 101; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
QY 1325 GTGCGGAGCGGCGTGTGCGCCAGGCGCAGCGGCGTGGAGAGCGCCAGGC 1384
DB 8 GTGGGCGGAGCGGTGGGCGCCGCGCCAGCGCCCGCCGCGCCCTCCGCTCCGCACTC 67
QY 1385 GGGCGCGCCCTCTTGCAGGAGCGCGCCAGCTCCCGCCCTGCGCTCCCTCC 1444
DB 68 GCGCTCTCCCTCCCTCCCGCCGCTCCCGCTCCCTCCCTCCCTCCAGCTGTC 127
QY 1445 GCTCGAGCTTACTTAACCTGCGCCGCGCGCGCGCGCTCTCACTTCCCTGAGCGCC 1504

